

Durham E-Theses

*Production of Recombinant Lectins from Garlic (*Allium Sativum*) Bulbs and Their Insecticidal Activity Against Hemipteran Insects.*

Wiles, Duncan Peter

How to cite:

Wiles, Duncan Peter (2008) *Production of Recombinant Lectins from Garlic (*Allium Sativum*) Bulbs and Their Insecticidal Activity Against Hemipteran Insects.*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/1324/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Academic Support Office, Durham University, University Office, Old Elvet, Durham DH1 3HP
e-mail: e-theses.admin@dur.ac.uk Tel: +44 0191 334 6107
<http://etheses.dur.ac.uk>

Production of Recombinant Lectins from Garlic (*Allium sativum*) Bulbs and Their Insecticidal Activity Against Hemipteran Insects

A thesis submitted by Duncan Peter Wiles B.Sc in accordance with the requirements of Durham University for the degree of Doctor of Philosophy.

The copyright of this thesis rests with the author or the university to which it was submitted. No quotation from it, or information derived from it may be published without the prior written consent of the author or university, and any information derived from it should be acknowledged.

Department of Biological and Biomedical Sciences

Durham University

June 2007



- 2 APR 2008

Abstract

Lectins have long been established as potential agents for use in control of insect pest species. In this study the use of garlic bulb lectins against phloem-feeding hemipteran insect pests is investigated. There is no *Bacillus thuringiensis* (Bt) toxin available which is effective against hemipteran insects. Plant lectins have been shown to be toxic to some insects, including aphids, in particular lectins from garlic. Subsequently heterodimeric (ASAI) and homodimeric (ASAI) garlic bulb lectins were cloned and expressed in a recombinant yeast system, *Pichia pastoris*. Recombinant lectins were successfully purified and demonstrated to be functionally active *in vitro* by haemagglutination assays and toxic to a range of hemipteran insects: pea aphid (*Acyrtosiphon pisum*), peach-potato aphid (*Myzus persicae*) and rice brown planthopper (*N. lugens*) in artificial diet bioassays.

To analyse the interactions between recombinant ASAI and the gut of *A. pisum* a pull-down assay was developed to identify specific interactions between ASAI and solubilized *A. pisum* gut proteins. Proteins which interacted with ASAI were subjected to matrix-assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometric analysis, this revealed that ASAI bound to alanyl aminopeptidase N (alanyl APN), a major constituent glycoprotein of the aphid gut which is rich in mannose oligosaccharides, and membrane-bound sucrase, an enzyme important in the maintenance of osmotic balance in the aphid gut. It was not possible to establish whether transport of ASAI into the haemolymph of *A. pisum* occurred using a Western blotting approach due to the lack of immunoreactivity of anti-ASA antibodies to the low levels of ASAI expected to be transported.

A fusion protein of ASAI-Avidin was created to enable the conjugation of biotinylated peptides of potential insecticidal interest to lectin carrier proteins. One such peptide, leucomyosuppressin (LMS) from cockroach *Diploptera punctata*, was conjugated to ASAI-Avidin and toxicity to *A. pisum* was demonstrated in artificial diet bioassays. It was confirmed that avidin is transported into the *A. pisum* haemolymph, suggesting that if ASAI was not responsible for transport of LMS in to the haemolymph then it may be mediated by avidin. Using a bioinformatic approach a putative *A. pisum* LMS gene was assembled.

Declaration

No material contained herein has been previously submitted for any other degree. Except where acknowledged, all material is the work of the author.

Statement of Copyright

The copyright of this thesis rests with the author. No quotation from it should be published without his prior written consent and information derived from it should be acknowledged.

Acknowledgments

For their support, advice and discussion I thank Drs. John Gatehouse and Elaine Fitches. Throughout this project their knowledge, experience and comments have been extremely valuable. Funding from Isagro Ricerca and the Yorkshire Agricultural Society is also gratefully acknowledged.

Some of the work described in this thesis was carried out in collaboration with other workers in the laboratory, the avidin gene and ASaII-avidin construct was prepared and produced by Dr. David Bown. The ASaII-RST constructs were assembled by Dr. Judith Philip. *Myzus persicae* bioassays were carried out and the data kindly supplied by Professor Angela Douglas. I would like to thank Prof. Douglas for advice on aphid bioassays and insect culture.

Thanks to all the members of Lab 1 past and present who have always ensured a pleasant working environment. I have made many friends throughout the course of this project. At some stage everyone has offered advice and support which is greatly appreciated. Thanks to Mr. John Gilroy for protein freeze-drying and Joanne Robson for assistance with MALDI-TOF.

I especially thank my parents who have always given support, guidance and encouragement throughout my education. Finally thank you to Kathryn whose constant support, encouragement and good sense have been absolutely invaluable to me.

Table of Contents

1	Introduction	1
1.1	What are Lectins?	1
1.2	The Evolution of Diversity in Carbohydrate-Binding Lectins	2
1.3	The Plant Lectins	5
1.4	The Mannose-Binding Lectins from <i>Allium Sativum</i> Bulbs	8
1.5	The Physiological Role of Plant Lectins	10
1.6	Lectin Binding to Insect Gut Proteins	12
1.7	Fusion Protein Technology	13
1.8	Hemipteran Insect Gut Physiology	15
1.8.1	The Insect Digestive Tract	15
1.8.2	The Hemipteran Digestive System	16
1.8.3	Proteolysis in Insects	16
1.8.4	Protein Absorption across the Insect Gut	18
2	Materials and Methods	21
2.1	Materials	21
2.1.1	Commonly Used Media	21
2.1.1.1	Yeast Culture Media	21
2.1.2	Commonly Used Buffers	21
2.1.2.1	DNA Electrophoresis	21
2.1.2.2	SDS-PAGE Buffers, Solutions and Size Markers	22
2.1.2.3	Western Blotting	23
2.1.2.4	Immunoblotting	24
2.1.2.5	ELISA Buffers	24
2.2	Standard Molecular Biological Techniques	25
2.2.1	Oligonucleotide Synthesis	25
2.2.2	Genomic DNA Extraction	25
2.2.3	Agarose Gel Electrophoresis	26

Table of Contents

vi

2.2.4	Purification of DNA from Agarose Gel	26
2.2.5	Bacterial Culture	26
2.2.6	Competent Cells	26
2.2.7	Transformation of <i>E. coli</i>	26
2.2.8	Isolation of Plasmid DNA	27
2.2.9	Restriction Endonuclease Digestion of DNA	27
2.2.10	Ethanol Precipitation of DNA	27
2.2.11	DNA Ligation	28
2.2.12	DNA Amplification using Polymerase Chain Reaction (PCR)	28
2.2.13	Ligation of PCR Products into Cloning Vector	28
2.2.14	Colony PCR	29
2.2.15	DNA Sequencing	29
2.2.16	Glycerol Stocks	29
2.3	Standard Biochemical Techniques	29
2.3.1	SDS-PAGE	29
2.3.2	Silver Staining	30
2.3.3	SYPRO Ruby Staining	30
2.3.4	Immunoblotting	31
2.3.5	Protein Concentration Determination	32
2.3.6	Deglycosylation of Protein Samples	32
2.3.7	Dialysis of Protein Samples	32
2.3.8	Lyophilization of Protein Samples	32
2.4	Production of Recombinant Proteins in <i>P. pastoris</i>	32
2.4.1	Transformation of <i>P. pastoris</i>	32
2.4.2	Expression Screening of Transformed <i>P. pastoris</i> Colonies	33
2.4.3	Overexpression of Recombinant Proteins in <i>P. pastoris</i>	33
2.4.4	Recovery of Protein extract from <i>P. pastoris</i>	34
2.5	Purification of Recombinant Garlic Lectins from <i>P. pastoris</i> Culture Supernatant	34
2.5.1	Purification of Recombinant Lectins by Hydrophobic-Interaction Chromatography	34
2.5.2	Purification of Recombinant Lectins by Anion-Exchange Chromatography	34

2.5.3	Purification of Recombinant Lectins by Cation-Exchange Chromatography	35
2.5.4	Purification of Recombinant Lectins by Metal-Affinity Chromatography	35
2.6	Haemagglutination Assays of Recombinant Proteins	34
2.7	Insect Culture	36
2.7.1	<i>N. lugens</i> stock culture	36
2.7.2	<i>N. lugens</i> artificial diet	36
2.7.3	<i>N. lugens</i> Artificial Diet Bioassay	37
2.7.4	<i>A. pisum</i> stock culture	39
2.7.5	<i>A. pisum</i> artificial diet	39
2.7.6	<i>A. pisum</i> Artificial Diet Bioassay	42
2.7.7	Insect Dissection and Haemolymph Extraction	42
2.8	Aphid Gut Pull-Down Assays	42
2.8.1	Preparation of ASaII-Magnetic Beads	42
2.8.2	Pull-Down Assay Sample Preparation	43
2.9	MALDI-TOF Analysis	43
2.9.1	Analysis of Recombinant Lectins by MALDI-TOF	44
2.9.2	Peptide Mass Fingerprint (PMF) Analysis of ASaII Binding Proteins	44
3	Production of Functional Recombinant ASaI & ASaII in <i>Pichia pastoris</i>	47
3.1	Introduction	47
3.2.1	Isolation of Full Length ASaI and ASaII Sequences from <i>Allium sativum</i> Genomic DNA	52
3.2.2	Preparation of Expression Constructs	53
3.2.3	Transformation of Expression Constructs into <i>Pichia pastoris</i>	58
3.2.4	Scale-up of Production of Recombinant Garlic Lectins in <i>Pichia pastoris</i>	62
3.3	Purification and Characterisation of Recombinant Garlic Lectins expressed in <i>P. pastoris</i> strain X-33	62
3.3.1	SDS-PAGE of recombinant ASaII produced in <i>P. pastoris</i> X-33	64
3.3.2	ASaI produced in <i>P. pastoris</i> strain X-33 shows variable polypeptide band patterns on SDS-PAGE	68
3.3.3	Analytical Gel Filtration to determine ASaII oligomerization	68

3.3.4	<i>in vitro</i> Haemagglutination Activity of Recombinant Garlic Lectins	69
3.3.5	Analysis of Recombinant Garlic Lectins by MALDI-TOF Spectroscopy	71
3.3.6	Glycosylation of recombinant ASAI produced in <i>P. pastoris</i> strain X-33	77
3.4	Purification and Characterisation of Recombinant Garlic Lectins expressed in <i>P. pastoris</i> strain SMD1168	81
3.4.1	SDS-PAGE of recombinant garlic lectin ASAI produced in <i>P. pastoris</i> SMD1168	81
3.4.2	Functional properties of recombinant garlic lectin ASAI produced in <i>P. pastoris</i> strain SMD1168	84
3.5	Discussion	86
4	Biological Activity of Recombinant Garlic Lectins on Insects	90
4.1	Introduction	90
4.2	Insect Artificial Diet Bioassays	91
4.2.1	Artificial Diet Bioassays Against Symbiotic and Aposymbiotic Pea Aphid (<i>Acyrtosiphon pisum</i>)	91
4.2.2	Artificial diet bioassays against symbiotic and aposymbiotic peach-potato aphid (<i>Myzus persicae</i>)	96
4.2.3	Artificial Diet Bioassays Against Rice Brown Planthopper (<i>N. lugens</i>)	100
4.3	Characterisation of Lectin Binding to Insect Gut	102
4.3.1	Isolation of Protein Fractions from <i>A. pisum</i>	102
4.3.2	Development of a Pull-Down Assay for Studying the Interaction of Recombinant Lectins with Aphid Gut Proteins	103
4.3.3	Pull-Down Assay of Whole <i>A. pisum</i> Insects	103
4.3.4	Pull-down assay of proteins bound by ASAI in the aphid gut	104
4.3.5	MALDI-TOF mass spectrometric analysis of aphid gut proteins bound by ASAI	107
4.4	Analysis of Lectin Transport into the Aphid Haemolymph	114
4.5	Discussion	116
5	Development of Insecticidal Fusion Proteins Based on Garlic Lectin	120

5.1	Introduction	120
5.1.1	Insecticidal fusion proteins containing a toxin and a lectin carrier	120
5.1.2	Insecticidal fusion proteins based on a carrier to which biotinylated peptides can be attached	121
5.2	ASAII-RST: A Recombinant Fusion Protein of Garlic Lectin and Red Scorpion Toxin	123
5.2.1	The ASAII ₁₂₇ -RST construct	123
5.2.2	The ASAII ₁₁₅ -RST construct	128
5.2.3	Artificial Diet Bioassay of ASAII ₁₁₅ -RST against <i>A.pisum</i>	128
5.3	The ASAII-Avidin fusion protein	133
5.3.1	Extraction of genomic DNA from chicken (<i>Gallus gallus</i>) embryo	133
5.3.2	Isolation of avidin coding sequence from chicken genomic DNA and assembly of avidin expression construct	133
5.3.3	Preparation of ASAII-Avidin expression construct	137
5.3.4	Production and Purification of ASAII-Avidin	140
5.3.5	SDS-PAGE analysis of recombinant ASAII-Avidin produced in <i>P. pastoris</i> SMD1168	141
5.3.6	ASAII-Avidin produced in <i>P. pastoris</i> strain SMD1168 shows variability in proteolytic degradation	146
5.4	ASAII-Avidin:Biotin-Leucomyosuppressin	149
5.4.1	Biotinylated Leucomyosuppressin (LMS) Production	149
5.4.2	Conjugation of ASAII-Avidin and Biotin-LMS	149
5.4.3	Artificial Diet Bioassays of ASAII-Avidin-LMS against <i>A.pisum</i>	152
5.4.4	<i>A.pisum</i> contains a gene encoding a putative LMS peptide hormone	153
5.5	Discussion	155
6	General Discussion	158
7	References	165

List of Tables

Table 2.1: Composition of *N. lugens* artificial diet MMD-1. 38

Table 2.2: Stock solutions for the formulation of minerals for inclusion in standard aphid diet. 39

Table 2.3 Stock solutions for the formulation vitamins for inclusion in standard aphid diet. 40

Table 2.4: Stock solutions for the formulation of amino acids for inclusion in standard aphid diet. 40

Table 3.1: Oligonucleotide sequences for ASAI and ASAII expression in pGAPzαA. 57

Table 3.2: Putative identification of mass ions present in MALDI-TOF mass spectrometric analysis of recombinant ASAI. 76

List of Figures

Figure 2.1: Feeding chamber used for <i>N. lugens</i> artificial diet bioassays	39
Figure 3.1: Agarose gel showing genomic DNA isolated from a single <i>A. sativum</i> bulb.	51
Figure 3.2: PCR of full-length ASAI and ASAII genes from <i>A. sativum</i> genomic DNA	51
Figure 3.3: ASAI full-length clone sequence	54
Figure 3.4: ASAII full-length clone sequence	56
Figure 3.5: ASAIpGAPzαA expression construct	59
Figure 3.6: ASAIIpGAPzαA expression construct	59
Figure 3.7: Sequences of predicted protein products from expression constructs for ASAII and ASAI.	60
Figure 3.8: Protein expression screen of ASAI in <i>P. pastoris</i> strain SMD1168	61
Figure 3.9: Protein expression screen of ASAII in <i>P. pastoris</i> strain SMD1168	61
Figure 3.10: Purification of recombinant ASAI (polypeptide 2), produced in wild-type <i>Pichia pastoris</i> (strain X-33)	65
Figure 3.11: Purification of recombinant garlic lectins from yeast culture supernatant by cation-exchange chromatography	66
Figure 3.12: Gel-filtration “Clean-up” purification profiles and SDS-PAGE analysis for ASAI and ASAII expressed in X33	67
Figure 3.13: Calibration of Superdex 200 column	70
Figure 3.14: Haemagglutination assays demonstrating <i>in vitro</i> biological activity of recombinant ASAI	70
Figure 3.15: MALDI-TOF mass spectrometric analysis of recombinant ASAII produced in <i>P. pastoris</i> strain X-33	72
Figure 3.16: MALDI-TOF mass spectrometric analysis of recombinant ASAI produced in <i>P. pastoris</i> strain X-33	75
Figure 3.17: SDS-PAGE showing analysis of partially proteolytically cleaved ASAI produced in <i>P. pastoris</i> by gel-filtration	78
Figure 3.18: Western blot of ASAI gel-filtration peak fractions	78
Figure 3.19: Deglycosylation of partially proteolytically cleaved ASAI produced in <i>P. pastoris</i> by treatment with N-glycosidase F	80

Figure 3.20: Purification of recombinant ASAI expressed in <i>P. pastoris</i> strain SMD1168	82
Figure 3.21: Gel-filtration “Clean-up” purification profiles and SDS-PAGE analysis for ASAI and ASAI expressed in SMD1168	83
Figure 3.22: SDS-PAGE showing mannan-agarose binding of recombinant garlic lectins expressed in <i>P. pastoris</i> strain SMD1168	85
Figure 4.1: Artificial diet bioassay of recombinant ASAI against symbiotic <i>A. pisum</i>	94
Figure 4.2: Artificial diet bioassay of recombinant ASAI against symbiotic <i>A. pisum</i>	94
Figure 4.3: Artificial diet bioassay of recombinant ASAI against aposymbiotic <i>A. pisum</i>	95
Figure 4.4: Artificial diet bioassay of recombinant ASAI against aposymbiotic <i>A. pisum</i>	95
Figure 4.5: Artificial diet bioassay of ASAI against symbiotic <i>Myzus persicae</i>	98
Figure 4.6: Artificial diet bioassay of ASAI against symbiotic <i>Myzus persicae</i>	98
Figure 4.7: Artificial diet bioassay of ASAI against aposymbiotic <i>Myzus persicae</i>	99
Figure 4.8: Artificial diet bioassay of ASAI against <i>Nilaparvata lugens</i>	101
Figure 4.9: Artificial diet bioassay of ASAI against <i>Nilaparvata lugens</i>	101
Figure 4.10: SDS-PAGE analysis of whole insect pull-down assay	105
Figure 4.11: SDS-PAGE analysis of insect gut tissue pull-down assay	106
Figure 4.12: SDS-PAGE of <i>Acyrtosiphon pisum</i> gut protein pull-down assay using SYPRO Ruby for MALDI-TOF analysis	109
Figure 4.13: MALDI-TOF spectrum of a 110 kDa band from an insect gut pull-down assay	110
Figure 4.14: Amino acid sequence data of membrane alanyl aminopeptidase N (ABD96614.1) from <i>A. pisum</i>	111

Figure 4.15: MALDI-TOF spectrum of a 66 kDa band from an insect gut pull-down assay	112
Figure 4.16: Predicted amino acid sequence of sucrase from <i>A. pisum</i>	113
Figure 4.17: Western blot analysis of ASaII-Avidin transport into the haemolymph of <i>A. pisum</i>	115
Figure 4.18: Western blot analysis of Avidin transport into the haemolymph of <i>A. pisum</i>	115
Figure 5.1: Sequence information for full-length ASaII ₁₂₇ -RST	126
Figure 5.2: SDS-PAGE analysis of ASaII ₁₂₇ -RST	127
Figure 5.3: Sequence information for truncated ASaII ₁₁₅ -RST	130
Figure 5.4: SDS-PAGE and Western blot analysis ASaII ₁₁₅ -RST	131
Figure 5.5: Bioassay of ASaII ₁₁₅ -RST fusion protein against <i>A. pisum</i>	132
Figure 5.6: Avidin gene sequence (<i>Gallus gallus</i> AVD gene; AJ311647) showing intron/exon structure	135
Figure 5.7: Primers and oligonucleotides used to assemble avidin coding sequence	136
Figure 5.8: Schematic diagram of avidin expression construct and predicted polypeptide sequence	136
Figure 5.9: Schematic diagram showing the assembly of the ASaII-Avidin expression construct	138
Figure 5.10: Sequence of the predicted polypeptide produced by the ASaII-avidin expression construct	139
Figure 5.11: Purification of ASaII-Avidin from yeast culture supernatant by ion-exchange chromatography	143
Figure 5.12: Purification of ASaII-Avidin expressed in <i>P. pastoris</i> strain SMD1168 by Ni ²⁺ affinity chromatography	144
Figure 5.13: “Clean-up” purification profile of ASaII-Avidin.	147
Figure 5.14: Intact ASaII-Avidin following affinity chromatography using Ni ²⁺ resin	148
Figure 5.15: Purification and verification of intact ASaII-Avidin:LMS	150
Figure 5.16: Artificial diet bioassay data for ASaII-Avidin:LMS fed <i>A. pisum</i>	151
Figure 5.17: Predicted sequence of <i>A. pisum</i> leucomyosuppressin prepro-peptide	

Abbreviations

Nucleic acid abbreviations

A – Adenine.

C – Cytosine.

G – Guanine.

T – Thymine.

Amino acid abbreviations

Name	Single letter code	Three letter code
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamic acid	E	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

-
- AaIT – *Androctonus australis* insecticidal toxin.
- AOX1 – Alcohol oxidase 1.
- APN – Aminopeptidase N.
- ASA – *Allium sativum* agglutinin.
- BPH – Brown planthopper, *Nilaparvata lugens*.
- BS³ – Bis[sulfosuccinimidyl] suberate.
- BSA – Bovine serum albumin.
- Bt* – *Bacillus thuringiensis*.
- Con A – Concanavalin A, *Canavalia ensiformis* agglutinin.
- ECL – Enhanced chemoluminescence.
- ELISA – Enzyme-linked immunoabsorbent assay.
- EST – Expressed sequence tag.
- FaRP – FMRFamide-related peptide.
- FITC – Fluorescein isothiocyanate-labelled albumin.
- FP – Fusion protein.
- GAP – Glyceraldehyde-3-phosphate.
- GAPDH – Glyceraldehyde-3-phosphate dehydrogenase.
- GFP – Green fluorescent protein.
- GNA – *Galanthus nivalis* agglutinin.
- LMS – Leucomyosuppressin.
- LRR – Leucine rich repeat.
- MALDI-TOF – Matrix-assisted laser desorption ionisation time of flight spectroscopy.
- ORF – Open reading frame.
- PBS – Phosphate buffered saline.
- PCR – Polymerase chain reaction.
- PHA – *Phaseolus vulgaris* agglutinin.
- PMF – Peptide mass fingerprinting/peptide mass fingerprint.
- PR – Pathogenesis related.
- RIPs – Ribosome-inactivating proteins.
- RST – Red scorpion toxin.

SDS-PAGE – Sodium dodecyl sulphate polyacrylamide gel electrophoresis.

TEMED – Tetramethylethylenediamine.

TFA – Trifluoroacetic acid.

TMOF – Trypsin modulating oostatic factor.

WGA – *Triticum aestivum* agglutinin.

1 Introduction

1.1 What are Lectins?

The term 'lectin' is derived from the latin *legere* (to pick out), to describe proteins isolated from plants with sugar-specific agglutination activity (Boyd, W. C., Shapleigh, E., 1954). Later discoveries of similar proteins in animal and bacterial systems have led to numerous attempts to define a lectin. The original definition includes many enzymes and immune system-derived proteins that are no longer considered to be lectins. In 1981 the Nomenclature Commission of the International Union of Biochemistry defined a lectin as a carbohydrate-binding protein of non-immune origin that agglutinates cells or precipitates polysaccharides or glycoconjugates. By limiting the definition to proteins of non-immune origin, anti-carbohydrate antibodies able to agglutinate cells are excluded from the group. According to this definition lectins must possess more than one carbohydrate-binding domain in order to have agglutination activity. This multivalency is essential for agglutination activity to be observed in accordance with such a definition. However, it was noted that this definition excluded poorly agglutinating proteins known to contain lectin subunits (eg: Ricin from *Ricinus communis*) and so the definition was extended to include these lectins.

More recently the definition has been further expanded after it was shown that some plant enzymes contain two subunits, a catalytic domain and a carbohydrate-binding domain such as in the Class I chitinases. An example of which is seen in *Hevea brasiliensis* (Beintema, J. J., 1994) where an N-terminal carbohydrate binding domain is in tandem with a C-terminal chitinase domain, these so called chimerolectins possess both specific binding activity and catalytic activity and are also seen in type 2 ribosome-inactivating proteins (RIPs) such as ricin, where a toxic N-terminal α -chain with typical N-glycosidase activity is linked to a C-terminal carbohydrate binding domain.

Other groups of lectins have also been classified according to their structural and functional similarities. In addition to the chimerolectins, the merolectin group has also been created after it was observed that some small proteins exhibit lectin-like binding to carbohydrates but

possess only a single binding domain. Previously such monovalent proteins were excluded from being classed as lectins because of their inability to agglutinate cells. Examples of merolectins include the chitin-binding hevein from *Hevea brasiliensis* and the mannose-binding lectin from *Epipactis helleborine* (Van Damme, E. J. M., *et al.*, 1994).

The majority of plant lectins that have been characterized to date fit into a class known as the hololectin class. These are proteins composed solely of carbohydrate-binding subunits containing more than one binding domain. The binding domains are identical or highly homologous in both structure and binding specificity. Lectins from *Galanthus nivalis* (GNA) (Van Damme, E. J. M., *et al.*, 1987), *Phaseolus vulgaris* (PHA) (Rigas, D. A., 1955), *Triticum aestivum* germ (WGA) (Raikhel, N. V., Wilkins, T. A., 1987) and *Allium sativum* (ASA) (Van Damme, E. J. M., *et al.*, 1992) are all hololectins. Accounting for the plant lectins being a broad group they have been defined by Peumans and Van Damme as proteins that possess at least one non-catalytic domain that binds reversibly to a specific carbohydrate (Powell, K. S., *et al.*, 1995).

1.2 The Evolution of Diversity in Carbohydrate-Binding Lectins

Oligosaccharides represent a group of molecules essential to all living organisms. Complex oligosaccharides have a wide variety of intracellular and extracellular roles; in particular carbohydrate groups are involved in intracellular and molecular 'recognition' roles. Extracellular oligosaccharides carry out structural roles such as acting as cell-surface ligands. The oligosaccharide component of the cell surface can dictate the interactions of the cell with other cells. In addition, oligosaccharides play essential roles in the formation of glycoconjugates within the cell, a process that is important for correct protein trafficking and targeting.

Lectins are proteins that are often involved in non-structural processes involving oligosaccharides; they are ubiquitous across all kingdoms of life from single-celled prokaryotes to higher mammals. Across this spectrum of organisms, carbohydrate binding and recognition processes occur in a wide variety of contexts, and subsequently the lectins

are a hugely diverse group of proteins. There are many different lectin families, most of which are evolutionarily unrelated, even where common structural folding and binding may occur, and no generalized architecture across these families is discernable from the wide range of lectin binding sites so far characterized. This has led to the belief that the wide range of carbohydrate binding specificities is the result of convergent evolution (Loris, R., 2002). The evolution of lectins is in truth more complex than this, with examples of both convergent and divergent evolution in both animal and plant lectins. Key to the evolution of a diverse range of lectins are two important structural aspects, protein folding and carbohydrate specificity. Overall protein structure is afforded by protein folding and it seems that in the case of lectins common protein folds are shared by many lectins but these may have evolved separately on different occasions, or may have derived from common ancestors in lower organisms. Similarly the exact evolutionary nature of many lectin carbohydrate-recognition domains (CRDs) is somewhat unclear; there is evidence of some CRDs being present in ancestral organisms and also evidence of divergence in CRDs in higher organisms leading to diversification of lectin function.

Despite commonly being complex and multidimeric proteins, the CRD of a lectin can often be attributed to a single module of the lectin polypeptide. In animals these CRDs are divided into groups according to their structure, the main families of animal lectins are the L-, C-, I- and P-type lectins (Drickamer, K., 1999, Fiedler, K., Simons, K., 1994), , (Angata, T., Brinkman-Van Der Linden, E. C. M., 2002), (Dahms, N. M., Hancock, M. K., 2002). In plants the classification of lectins is confined to seven families according to the specificity of their CRDs, four major families: the legume lectins, chitin-binding lectins, Type 2 RIPs and monocot mannose-binding lectins and 3 smaller groups: the jacalin-related lectins, Cucurbitaceae phloem lectins and the amaranthin group. These seven groups account for virtually every plant lectin (Van Damme, E. J., *et al.*, 1998). Plant lectins are discussed in section 1.3.

The diversity of lectins is great but as well as functional differences across kingdoms, some structural and sequence similarities can be observed. Several higher plant lectins belong to protein families of prokaryotic origin such as lectins possessing ricin B domains, or

Galanthus nivalis agglutinin (GNA) domains which have some sequence similarity with proteins in bacteria as well as non-plant eukaryotes (Van Damme, E. J., *et al.*, 2007). Similarly some animal lectins have obvious ancestors in lower unicellular and multicellular eukaryotes, for example the gene sequence of human calnexin, a lectin involved in the retention of incorrectly folded glycoproteins in the endoplasmic reticulum (ER), shares 43% homology with calnexin from the single-celled yeast *Schizosaccharomyces pombe* (Dodd, R. B., Drickamer, K., 2001). Other lectins have less obvious ancestral origins in lower organisms but still show relatedness to non-lectin protein domains. For example the hevein domain which makes up part of the chitin-binding lectins found in plants is also widespread in fungi, indicating that this CRD was present in an earlier eukaryotic ancestor.

The legume lectin domain is classified in the same protein superfamily as the vesicular integral membrane protein 36 (VIP36) found in animals and fungi, and the endoplasmic reticulum-Golgi-intermediate compartment 53-kDa protein (ERGIC-53), examples of L-type lectins. The β -sandwich fold seen in the legume lectins is also present in animal galectins and pentraxins, further members of the group known as the L-type lectins, as well as some carbohydrate-processing and other enzymes such as β -glucanase and asparagine amidase (Kuhn, P., *et al.*, 1994), (Keitel, U., *et al.*, 1993). The leguminous plant lectins are soluble, secreted proteins and, like most plant lectins are highly expressed in specialized tissues (often most abundantly in storage tissues such as bulbs), whereas the animal L-type lectins are low abundance, membrane-bound luminal proteins found in many different cell types. The work of Dodd and Drickamer (2001) suggests that the different L-type lectins retain similar mechanisms of sugar binding and that the L-type CRD fold is ancient, as a related homologous ancestor exists in *S. cerevisiae*. However, Loris (2002) states that the positioning of the CRDs on the β -sandwich structure is different in plant and animal lectins and that no similarity in CRD architecture exists. It therefore seems uncertain exactly how the legume lectin-like β -sandwich fold evolved. What is evident is that despite a similar structural fold, these lectins carry out diverse functions across different kingdoms.

1.3 The Plant Lectins

An understanding of the structural and evolutionary inter-relatedness of plant lectins has been greatly aided by the development of recombinant DNA technologies, enabling cloning of a large number of plant lectin genes across a wide range of plant species. Subsequently it has been possible to divide the heterogenous plant lectins into the seven groups described above, the four main groups: Legume lectins, chitin-binding lectins, Type 2 RIPs and monocot mannose-binding lectins and the further 3 small families: the jacalin-related lectins, Cucurbitaceae phloem lectins and the amaranthin group.

The legume lectins are a large family of proteins from the plant family Leguminosae and to date over 70 lectins have been identified from a range of species within this group. As with most lectins, legume lectins accumulate predominantly in storage tissues, most of the lectins isolated have been taken from mature seeds, where lectins have been shown to account for up to 10% of total soluble protein. A wide range of carbohydrate-binding specificities occur across the legume lectins, the most common group being the mannose/glucose-binding lectins which includes *Canavalia ensiformis* lectin (Con A) and *Pisum sativum* lectin and the galactose/N-acetylgalactosamine specific lectins from peanut (*Arachis hypogaea*) and pagoda tree (*Sophora japonica*). Some legume lectins do not show any carbohydrate-binding specificity to simple monosaccharides, such as kidney bean lectin (*Phaseolus vulgaris*), which interacts solely with oligosaccharides. Legume lectins, like other lectin groups are synthesized as preprolectins from which initially the signal peptide is cleaved followed by the removal of a C-terminal peptide. Following this initial modification some legume lectins undergo a further excision of an internal peptide sequence. Lectins that undergo internal peptide cleavage have been termed 'two-chain' legume lectins; those which do not are termed 'one-chain' legume lectins. The site of the internal cleavage of the two-chain legume lectins vary according to species, in some instances similar sized polypeptides are produced such as the mannose-binding bark lectin from *S. japonica* or different sized as seen in the *P. sativum* lectin. An exception to the above post-translational processing has been demonstrated for Con A. Mature Con A is formed from an initial glycosylated precursor, which is first deglycosylated and then cleaved into smaller peptides which are reannealed to

give an intact subunit in which amino acids 1-118 and 119-237 were reversed in the final lectin form (Bowles, D. J., *et al.*, 1986).

The chitin-binding lectins possess one or more hevein domains. The hevein domain is a small chitin-binding protein from the rubber tree (*Hevea brasiliensis*) and is 43 amino acids in length. Full-length hevein is a larger precursor of 144 amino acids but only the 43 amino acid domain is present in chitin-binding lectins. This active chitin-binding domain is present in the chitin-binding lectins from many taxonomically unrelated plant families such as Gramineae, Solanaceae and Euphorbiaceae. Most chitin-binding lectins are in the hololectin form, although monovalent merolectins also exist as do the Class I chitinases, a group of chimerolectins containing a chitin-binding domain attached to a catalytic domain with chitinase activity. The chitin-binding lectins from the Gramineae have been most widely studied and lectins have been isolated from a number of cereal species, in particular *Triticum aestivum*, *Secale cereale*, *Hordeum vulgare* and *Oryza sativa* (Peumans, W. J., Stinissen, H. M., 1983). All these lectins are 18 kDa dimeric, unglycosylated proteins; typically of the chitin-binding lectins they are rich in glycine and cysteine residues. The lectin from rice (*O. sativa*) differs in that it is composed of an 8 kDa and a 10 kDa polypeptide, products of proteolytic cleavage of the 18 kDa precursor. Tetraploid and hexaploid cereal species contain two or three highly homologous isoelectin genes respectively, whereas diploid species contain just a single lectin gene. Like the legume lectins, gramineae lectins are synthesised on the ER and are translated as precursors that undergo C-terminal cleavage of a glycosylated peptide to form mature polypeptides. Wheat germ agglutinin (WGA) has been crystallized and it has been shown that each of the four hevein domains is stabilized by the presence of a disulphide bridge.

Type 2 RIPs are a well understood group of chimerolectins. Type 2 RIPs are typically made up of one, two or four identical units, each unit containing two structurally distinct polypeptide subunits called the A and B chains. Similarly to the class I chitinases, the A chain of the type 2 RIPs has an enzymatic function, in this case it is a highly specific rRNA N-glycosidase activity, which is responsible for the inactivation of eukaryotic ribosomes. The B chain possesses a carbohydrate-binding activity comparable to a lectin. The carbohydrate-

binding specificity of Type 2 RIPs is mostly for galactose or N-acetylgalactosamine, although other specificities exist in some species. The first sequence data obtained for Type 2 RIPs was taken from *Abrus precatorius* seeds (arbrin) (Wood, K. A., *et al.*, 1991), *Ricinus communis* seeds (ricin) (Lamb, F. I., *et al.*, 1985), *Sambucus nigra* bark (SNA-I and SNA-V) (Van Damme, E. J., *et al.*, 1996b), (Van Damme, E. J., *et al.*, 1996a) and *Sambucus sieboldiana* bark (SSA) (Kaku, H., *et al.*, 1996). A high degree of sequence similarity was observed between these genes, suggesting that these lectins form part of a superfamily of structurally-related proteins across taxonomically unrelated families. The A and B chains of Type 2 RIPs are always synthesized as larger precursors joined by a linker region linking the C-terminal of the first domain to the N-terminal of the second domain. Once the linker region is excised the two domains remain linked by a disulphide bridge. The three-dimensional structure of the most well studied of the Type 2 RIPs, ricin, has been resolved by X-ray crystallography and has shown that both the A and B chains are globular structures, the A chain having extensive secondary structure and the B chain having two globular domains, each of which is able to bind lactose in a shallow cleft. Both of these B chain domains are similar and are thought to have arisen through gene duplication (Rutenber, E., Robertus, J. D., 1991). Like the chitin-binding lectin WGA, the stability of ricin has been shown to be dependent on disulphide bridges.

The first plant lectin to be identified as a mannose-binding lectin was isolated from snowdrop (*Galanthus nivalis*) (Van Damme, E. J., *et al.*, 1987). Subsequently mannose-specific lectins have been found in the Amaryllidaceae, Alliaceae, Araceae, Orchidaceae and Liliaceae. As all of these families represent monocotyledenous plants the group of lectins was termed the 'monocot mannose-binding lectins'. Like the other groups of lectins the corresponding gene sequences reveal a superfamily of evolutionarily related proteins with significant sequence homology. Despite the sequence similarity and the fact that the core polypeptide unit of the monocot mannose-binding lectins is always approximately 12 kDa, important differences in post-translational modification and processing occur between specific lectins. Most commonly monocot mannose-binding lectins are homomeric proteins of two or four identical subunits, synthesized as distinct polypeptides. Like most plant lectins from all families, these polypeptides are synthesized as lectin precursors that undergo signal peptide cleavage and

removal of a C-terminal peptide sequence (which in the case of the monocot mannose-binding lectins is usually non-glycosylated, in contrast to the Type 2 RIPs). In contrast to all of the other groups of plant lectins described so far the monocot mannose-binding lectins are encoded by closely related gene families rather than by single genes, as seen in the case of *Galanthus* and the lectins from garlic bulbs (Van Damme, E. J., *et al.*, 1991b), (Van Damme, E. J., *et al.*, 1992). The structure of *Galanthus nivalis* agglutinin (GNA) has been resolved by X-ray crystallography and showed a β -barrel of three antiparallel four-stranded β -sheets joined by loops. Each monomeric unit contains three identical binding sites with 4 conserved amino acid residues. The structure of other monocot mannose-binding lectins is postulated to resemble that of GNA given the high degree of sequence similarity between the different lectins. Using GNA as a three-dimensional template for modelling other monocot mannose-binding lectins it has been shown that the structural resemblance is strong. One significant difference between the monocot mannose-binding lectins is their valency, GNA contains three binding sites per monomer making a total of twelve in the fully formed tetramer, bulb lectins from garlic (*Allium sativum*) are dimeric and therefore only possess six binding sites per lectin polypeptide, this could explain their reduced affinity towards mannose compared to GNA.

1.4 The Mannose-Binding Lectins from *Allium Sativum* Bulbs

Mannose-binding lectins are present in the leaves, roots and bulbs of garlic. Four garlic lectins have been isolated, two from bulbs (ASAI and ASAI), one from leaves (ASAL) and one from roots (ASAR). Initially polypeptides of 11500 Da and 12500 Da were isolated from garlic bulbs by mannose affinity chromatography and shown to form dimers by gel-filtration chromatography (Van Damme, E. J., *et al.*, 1991a). It was later discovered two bulb lectin isoforms existed in garlic bulbs (Van Damme, E. J., *et al.*, 1992), the genes were identified from screening of a cDNA library using amino acid sequences obtained from protein sequencing (Van Damme, E. J., *et al.*, 1991a). The two different lectins became known as ASAI, the heterodimeric lectin from garlic bulb lectin and ASAI, the homodimeric lectin. The genes were identified from a garlic bulb cDNA library using degenerate primers based upon amino acid sequence data gained from the work of Kaku *et al.* (Kaku, H., *et al.*, 1992).

The molecular structures of the garlic lectins differ from those of the Amaryllidaceae family, which includes GNA. Whereas Amaryllidaceae lectins are composed of identical 12.5kDa subunits, usually in tetrameric form, garlic lectins exist in two forms: as heterodimers of an 11.5kDa subunit and a 12.5kDa subunit (ASAI), and as homodimers of a 12kDa subunit (ASAI). These two lectin isoforms have been shown to be encoded by two different genes (Van Damme, E. J., *et al.*, 1992). ASAI is translated from an mRNA of approximately 1400 nucleotides, containing a 940bp open reading frame (ORF) encoding 306 amino acids, the first twenty-eight of which represent signal peptide, and are cleaved after translation. The calculated molecular mass of this 306 amino acid polypeptide is 33020 Da. As well as two polypeptide subunits of 11500 Da and 12500 Da which form the mature lectin subunits, ASAI also contains a linker region between the two subunits which contains an N-glycosylation site. Gel-filtration of native ASAI from garlic bulb extract showed a single protein elute at approximately 25000 Da, implying that the two subunits run as a dimer. The predicted size of the 306 amino acid polypeptide and the sizes of ASAI subunits observed by SDS-PAGE and the results of gel filtration all differ significantly suggesting that a number of processing steps of the initial lectin precursor are taking place. N-terminal sequencing has shown that ASAI loses a 28 amino acid signal peptide and that the first, 11.5 kDa subunit of ASAI begins at residue 37 (RNLL...) and the second, 12.5kDa begins at residue 184 (RNIL...). This indicated that both subunits were produced from a single precursor (Van Damme, E. J., *et al.*, 1992). ASAI is translated from an mRNA of approximately 800 nucleotides, encoding a polypeptide of 154 amino acids. The coding region of ASAI is almost identical to that of the 12.5kDa subunit of ASAI. The processing of the ASAI homodimer is thought to be similar to that of the Amaryllidaceae lectins (Van Damme, E. J., *et al.*, 1992). There is some level of diversity between lectin coding sequences from different garlic bulbs although overall similarity is over 95%, which is greater than that of the Amaryllidaceae lectins.

The structure of garlic lectins and other mannose-specific lectins has been elucidated and the structural stability of ASAI has also been studied in detail. The subunits possess a unique interrelated pseudo-3-fold-symmetry having three 4-stranded antiparallel sheets orientated as

3 sides of a trigonal prism forming a 12-stranded barrel. The core of this barrel is lined with hydrophobic residues, which stabilize the fold. Dimerization of ASAI occurs by strand exchange between three strands of the first subunit and one of the second. The unfolding of ASAI has been shown to be a two-phase, reversible process implicating quaternary interactions as important in conformational stability (Bachhawat, K., *et al.*, 2001).

The sugar specificity of garlic bulb lectins has been studied and compared to other mannose-binding lectins, ASA binds single mannose residues weakly. Invertase has been demonstrated to be the strongest glycoprotein ligand. The reasons for this are thought to be the large number of α 1-2-linked mannose residues present in invertase. The presence of α 1-2-linked mannose residues at the non-reducing ends of high mannose oligosaccharides has been shown to increase ASA binding (Dam, T. K., *et al.*, 1998). GNA recognizes only terminal α 1-3-linked mannose residues. This suggests that the mannose binding affinity of ASA to the insect gut will differ from that of GNA according to the structure of the mannose oligosaccharides present there.

1.5 The Physiological Role of Plant Lectins

Lectins are found predominantly as storage proteins *in planta*. Commonly plant lectins are found in abundance in seeds and bulbs, although they may also be found in the leaves of some plants. The role of lectins within the plant remains unclear. The specific carbohydrate binding of lectins has been used to suggest that some lectins may play an important role in specific recognition processes *in planta*, although very few receptors with specific carbohydrate moieties to enable lectin binding have been identified in plants. This raised the possibility that perhaps the specific receptors for lectin binding were to be found outside of the plant itself. This idea, combined with the knowledge that lectins are found predominantly in metabolically inactive storage tissues led to the suggestion that the primary role of lectins may lay in passive plant defence.

The interactions of lectins with foreign organisms have been studied in detail. Perhaps the best understood example being that of the type 2 RIP ricin. Ricin is acutely toxic to all eukaryotes including humans. Its mode of action has been determined and it illustrates clearly

how lectins can act as defence agents. Consisting of 2 domains, an A chain which contains a catalytic domain, which cleaves the *N*-glycosidic bond of an adenosine residue of rRNA. The B chain of ricin is the carbohydrate-binding domain, which binds to receptors on the cell surface. Ricin is however a chimerolectin, hololectins are far more commonly occurring in plants and do not contain catalytic domains.

Although it is unlikely that all plant lectins play a role in plant defence, lectins that are present at high concentrations in the plant tissue have been shown to play a role in plant defence. The ability of lectins to bind glycoconjugates from other organisms provides a strong suggestion that lectins are involved in the plant's defence strategy. Lectins can bind simple sugars such as mannose, galactose or glucose, as well as simple disaccharides such as lactose and maltose; however many lectins have greater affinity to longer oligosaccharides that are rare in plants but more common in animals and fungi. One of the main groups of plant lectins are the chitin-binding lectins, chitin is absent from plants but is a major constituent the fungal cell wall as well as the cytoskeleton of insects. Another example is the sialic acid-binding lectin from *Maackia amurensis* (Knibbs, R., *et al.*, 1991). Sialic acid is a nine-carbon monosaccharide that is a common component of animal and bacterial glycoproteins, which is absent from plants. Such observations along with the lack of lectin receptors identified in plants firmly suggest a defensive role for most lectins. Furthermore, lectins are extremely stable and resistant to proteolysis by both digestive enzymes and gut bacteria; subsequently lectins retain full activity in the gut of insects of higher animals. A number of plant lectins have been shown to have negative effects on insect survival when administered through artificial diets and transgenic plants; for example GNA induced a 42% nymphal mortality on the peach potato aphid (*Myzus persicae*) at 1500 µg/ml in artificial diet (Sauvion, N., *et al.*, 1996) and transgenic maize plants expressing GNA at over 0.22% total soluble protein showed enhanced resistance to corn leaf aphid (*Rhopalosiphum maidis* Fitch) under greenhouse conditions where nymph production was reduced by 46.9% compared to control plants (Wang, Z. Y., *et al.*, 2005). Sauvion *et al.* demonstrated that Con A produced a feeding deterrence at 400 µg/ml when fed to pea aphid (*Acyrtosiphon pisum*) and that this deterrence was a result of intoxication and not a sensory mediated process (Sauvion, N., *et al.*, 2004). The insecticidal effects of garlic leaf lectins have been described against the hemipteran species red cotton bug

(*Dysdercus cingulatus*) and mustard aphid (*Lipaphis erysimi*) in artificial diet bioassays (Bandyopadhyay, S., *et al.*, 2001), in transgenic tobacco and transgenic rice plants where the survival of rice brown plant hopper and green leaf hopper were reduced to 36% and 40.5% respectively compared to control plants (Dutta, I., *et al.*, 2005, Saha, P., *et al.*, 2006).

1.6 Lectin Binding to Insect Gut Proteins

The binding of plant lectins to carbohydrate residues in the insect gut has been well studied. The scope for lectin binding in the insect gut is vast; many membrane proteins including hormone receptors, transport proteins and brush-border enzymes are glycosylated in the brush-border membrane providing potential binding sites for lectins. Lectins have also been shown to bind bacterial endosymbionts in the insect gut (Banerjee, S. *et al.*, 2004) (Foissac, X., *et al.*, 2002). The effects of the mannose-binding family of lectins on insects have been investigated in the greatest depth. Mannose is a sugar found commonly in the midgut of insects but rarely in higher animals. GNA is the most studied mannose binding lectin, however, there is a large amount of similarity between mannose binding lectin DNA sequences, and GNA may not show the strongest affinity for binding to the midgut of most insects.

The toxin from *Bacillus thuringiensis* (Bt) has been demonstrated to bind a mannose-rich glycoprotein enzyme, aminopeptidase, in the midgut membrane of several lepidopteran species including tobacco hornworm (*Manduca sexta*) (Knight, P. J., *et al.*, 1995), tobacco budworm (*Heliothis virescens*) (Gill, S. S. *et al.*, 1995), silkworm (*Bombyx mori*) (Yaoi, K., *et al.* 1999). An aminopeptidase has recently been characterized in the pea aphid, *A. pisum* (Cristofolletti, P. T., *et al.*, 2006); this protein represents over 15% of total midgut protein in the pea aphid and, like other aminopeptidase enzymes, is rich in mannose oligosaccharides. Such highly glycosylated proteins represent potential binding targets for mannose-binding lectins in the insect midgut.

GNA has been shown to bind a 26kDa subunit of a ferritin-like glycoprotein in the midgut of the hemipteran pest rice brown planthopper *Nilaparvata lugens* (Du, J., *et al.*, 2000). The

ability of lectins to be translocated across the midgut and into the insect haemolymph has been demonstrated. The L- isoform of the lectin from *Phaseolus vulgaris* (PHA-L), GNA and the lectin from *Canavalia ensiformis* (Con A) have all been detected in the haemolymph of tomato moth (*Lacanobia oleracea*) larvae (Fitches, E. C. *et al.*, 2001), (Fitches, E. C. *et al.*, 2004). Despite these findings it has become clear from these findings that insect pests cannot be fully controlled by using lectins alone as lectins constitute a passive plant defence strategy rather than a high level of acute toxicity. Consequently a new strategy is required that can produce higher levels of oral toxicity, the ability of lectins to cross the insect gut lumen intact makes them a candidate ‘carrier’ protein for use in coupling to molecules of high toxicity to insects such as toxins, venoms and hormones in a ‘fusion-protein’.

1.7 Fusion-Protein Technology

The fusion protein approach takes advantage of lectins as carriers utilizing their ability to cross the insect gut into the haemolymph. Using recombinant DNA technology allows the fusion of potentially insecticidal proteins and peptides to lectin carriers. This approach relies on the lectin to deliver the insecticidal subunit of the fusion protein across the insect gut into the haemolymph in an active form. Using the recombinant protein approach it has become possible to produce and evaluate a wide range of potential proteins and peptides for their insecticidal properties. Typically fusion proteins are used to deliver toxic proteins or peptides which are ineffective when administered orally due to breakdown in the insect gut by digestive enzymes. Injection assays can be used to demonstrate potential toxicity of such toxins when delivered to the insect haemolymph. The production of a recombinant fusion protein containing a lectin carrier and insecticidal toxin provides a novel mode of action, allowing the toxin subunit to be active when orally administered. To date four fusion proteins have been produced, in each case the lectin carrier subunit was derived from GNA. The first fusion protein (FP1) incorporated an allatostatin from *M. sexta* (Manse-AS), a peptide hormone which is involved in the control of gut peristalsis and may inhibit the production of juvenile hormones. This fusion protein was produced in *E. coli* and when fed to tomato moth larvae (*Lacanobia oleracea*) a decrease in larval weight was observed corresponding to a significant reduction in artificial diet consumption compared to control, GNA and Manse-AS

fed insects and increased levels of allatostatin could be immunodetected in the haemolymph of FP1 fed insects compared to Manse-AS fed insects (Fitches, E., *et al.*, 2002). The second fusion protein (FP2) was based around a neurotoxin venom peptide from the arthropod *Segestria florentina* (SFI1). This fusion protein was produced successfully in the yeast *Pichia pastoris*. Artificial diet bioassays of the purified recombinant fusion protein demonstrated that neither GNA, nor SFI1 was toxic to *L. oleracea*, but FP2 fed at 2.5% of dietary protein resulted in 100% mortality of first stadium larvae after 6 days (Fitches, E., *et al.*, 2004a). The ability of lectins to act as carriers to large toxic proteins would make many enzymes amenable as constituents of fusion proteins. The viability of such an enzyme in fusion proteins was investigated by the development of recombinant protein containing GNA with chitinase from *L. oleracea* (~65 kDa) linked at its N-terminus (FP3) (Fitches, E., *et al.*, 2004b). Chitin is an integral part of the peritrophic membrane (PM) found in the gut of most insect species and chitinase expression has been shown to be developmentally and hormonally regulated to coincide with moulting and metamorphosis in lepidopteran larvae. The delivery of exogenous chitinase into the insect may disrupt the endogenously controlled turnover of the insect's PM. Recombinant chitinase and FP3 were both produced in *P. pastoris* and in injection assays both were both highly toxic to *L. oleracea*, producing 100% mortality in hours at concentrations as low as 4 µg/g insect tissue. Microscopic analysis showed that chitinase-fed insects had significantly narrower endo-cuticle and exo-cuticle chitin layers, an effect which was not observed in FP3 fed insects. In bioassays recombinant chitinase and FP3 had similar effects on insect weight gain and diet consumption, producing 60% and 56% less weight gain respectively, compared to control insects. This corresponded to a diet consumption of 66% and 60% of that of control insects for recombinant chitinase and FP3 respectively. Western blotting analysis showed that no chitinase was present in the haemolymph of either recombinant chitinase or FP3 fed insects. GNA was present in the haemolymph of FP3 fed insects suggesting that proteolytic cleavage of the fusion protein had occurred in the gut of the insect. In this case no benefit was gained from fusing chitinase to a lectin carrier, demonstrating that transport of chitinase to the insect haemolymph is not necessary for an insecticidal function and that no novel activity is obtained from the fusion protein. The use of venom proteins as constituents of fusion proteins was investigated further using a gene encoding a toxin (RST) from the red scorpion (*Mesobuthus tamulus*). Upon

injection the GNA-RST fusion protein (FP4) reduced survival of *L. oleracea* to 30% after 7 days at doses in the order of 1-10 µg per insect, a dose of 20 µg per insect reduced survival to 40% after 48 hours. At as much as 4.5% total dietary protein little acute toxicity was observed against third stadium *L. oleracea* larvae, survival was reduced to 75% after 12 days. Intact and cleaved FP4 was observed in the haemolymph of dissected insects, demonstrating that although some proteolytic cleavage had occurred, some intact fusion protein was present in the haemolymph. Little research has been carried out on potential alternative carriers to GNA, such as garlic lectins. Many factors may affect the success of lectins as carriers such as carbohydrate-binding specificity; a lectin may be chosen for specific insect species where certain oligosaccharides that predominate in glycoproteins exhibit the greatest binding affinity for the carrier. The success of a carrier may also be related to its valency, GNA forms tetrameric units, whereas garlic lectins form dimers, in the context of fusion proteins this may have implications on fusion protein stability and the ability of the fusion protein to bind to gut receptors and in turn cross into the haemolymph.

1.8 Hemipteran Insect Gut Physiology

1.8.1 The Insect Digestive Tract

In insects excretion occurs via the Malpighian tubules and the hindgut. The Malpighian tubules are located at the midgut-hindgut junction and lie freely within the haemolymph of the main body cavity. The role of the Malpighian tubules in insects is analogous to that of the nephron in vertebrates. In insects however, primary urine is formed by the active transport of KCl or NaCl into the tubule lumen producing osmotic filtration rather than ultrafiltration as seen in the vertebrate nephron. This active transport process is essential to the insect as it has been demonstrated to facilitate the removal of toxic plant alkaloids and presumably other large plant secondary metabolites from the Malpighian tubules (Madrell, S. H. P., Gardiner, B. O. C., 1976). However, in hemipteran insects the physiology of the gut differs from that of most other insects including the so-called 'chewing' insects. The hemipteran gut contains a convoluted region called the filter chamber. In the filter chamber the gut is folded such that the initial midgut region lies against the terminal midgut region. This creates an osmotic

gradient across the epithelial cells of the initial and terminal midgut regions, allowing excess dietary fluid to pass to the terminal midgut and proximal Malpighian tubules bypassing the majority of the insect midgut. This highly differentiated section of digestive tract enables hemipteran pests to ingest large amounts of dietary fluid.

1.8.2 The Hemipteran Digestive System

Insects of the order Hemiptera possess characteristics that are thought to be adaptations acquired to enable feeding by phloem extraction. Unlike phytophagous insects hemipteran insects do not possess the crop region of the foregut. The crop is involved in storage of ingested food prior to digestion and is seen in all other insect orders. Hemipteran insects also lack many of the polymer and oligomer hydrolases that are commonly seen in other orders. Enzymes such as amylase, cellulase and hemicellulase involved in the hydrolysis of starch and cellulose polymers are essential to phytophagous insects but are not found in hemipteran insects. Similarly the peritrophic membrane is absent from hemipteran insects, the peritrophic membrane is an important physiological feature of all other insects that provides mechanical protection for midgut cells from large food particles. The peritrophic membrane is a chitinous protein-carbohydrate matrix which develops within the gut itself. It is supposed that the absence of the peritrophic membrane in hemipteran insects is associated with the absence of significant levels of midgut luminal digestion. Hemipteran insects possess a perimicrovillar membrane which maintains a constant distance from the microvillar membrane. The perimicrovillar membrane is bereft of intramembranous particles and allows potassium ions from the perimicrovillar space to be actively transported across the microvillar membrane into midgut cells producing a concentration gradient between the perimicrovillar space and the gut used for active absorption of organic compounds into the midgut cells.

1.8.3 Proteolysis in Insects

For the successful deployment of lectins targeted against hemipteran insect pests, knowledge of their specialist feeding behaviour is required. The feeding of phloem-extracting insects differs from both lepidopteran and coleopteran pests. In lepidopteran and coleopteran insects hydrolysis of dietary protein is essential for the procurement of sufficient nitrogen levels. The

importance of proteolytic enzymes to these orders of insect pests is reflected in the presence of plant proteinase inhibitors as a central part of the plant's defence against them. Hemipteran insects have been considered in the past to completely lack proteolytic activity in the midgut. Studies on the aphid midgut (Rahbe, Y., *et al.*, 1995), and salivary exudates (Cherqui, A., Tjallingii, W. F., 2000) found no evidence of proteinase activity which is suggested to be a consequence of hemipteran insects' specialized feeding habits. By feeding on the xylem and phloem sap, hemipteran insects are believed to obtain nitrogen from reduced nitrogen compounds in the vascular tissues, chiefly in the form of glutamine and asparagine as well as other free amino acids. The requirement for proteinase activity was thought to be minimal. Some proteins, however have been found in the phloem sap, for example thioredoxin in the phloem of rice plants (Ishiwatari, Y., *et al.*, 1995) and subsequently proteinase activity has been detected in the rice brown plant hopper, *Nilaparvata lugens* (BPH) (Foissac, X., *et al.*, 2002). However, BPH still obtains most of its nitrogen in the form of free amino acids. BPH also feeds on sucrose of up to 1 M concentration in the phloem sap; this exerts osmotic pressure of up to three times that of the insect's body fluids. Consequently, maintenance of water balance in BPH is crucial for the insect to survive.

In the pea aphid (*Acyrtosiphon pisum*) a similar situation is present. *A. pisum* is a phloem feeding insect which encounters high levels of sucrose (0.15-0.73 M) and some amino acids (15-65 mM) in the host plant phloem sap. High levels of digestive activity in the aphid gut have been traditionally been regarded as unnecessary despite the presence of some proteins in the phloem sap. The most important functions of the aphid midgut are the hydrolysis of sucrose and the putative absorption of free amino acids. The importance of sucrose hydrolysis is critical to the insect; sucrose is unable to be assimilated across the insect gut presenting osmotic and nutritional problems. The hydrolysis of sucrose is carried out by a single α -glucosidase enzyme termed 'sucrase' which is localized to the distal midgut (Price, D. R., *et al.*, 2007). This enzyme converts sucrose to its constituents: glucose and fructose.

1.8.4 Protein Absorption Across the Insect Gut

It has been known for some time that peptides and proteins are able to cross the insect gut into the haemolymph, but the mechanism by which this takes place has remained unclear. The transepithelial transport of several peptides and proteins have been demonstrated, for instance the small trypsin-modifying oostatic factor (TMOF) in *Aedes aegypti* (Borovsky, D., Mahmood, F., 1995) and the small venom protein from *Androctonus australis* (AaIT) being transported across the midgut of the flesh fly *Sarcophaga falculata*; transport of larger proteins has also been demonstrated, for example horseradish peroxidase (Fishman, L., Zlotkin, E., 1984) and green fluorescent protein (GFP) (Habibi, J., *et al.*, 2002). Horseradish peroxidase has been demonstrated to cross the insect midgut along intercellular junctions (Modespacher, U. P., *et al.*, 1986) and Habibi *et al.* (2002) demonstrated that GFP crosses the midgut through a paracellular route in the hemipteran insect *Lygus hesperus*. Transport of proteins across the insect gut through septate junctions has also been suggested as a mechanism for transepithelial movement, although the exact nature of the epithelium and its permeability has been shown to vary between insect species. Septate junctions resemble mammalian tight junctions and are responsible for controlling the permeability of this type of paracellular pathway. Septate junctions are leaky and studies have shown that molecules as large as 5 kDa are able to pass through them (Zhu, W., *et al.*, 2001). The work of Casartelli *et al.* demonstrated that albumin was transported across the midgut of *Bombyx mori* intact (Casartelli, M., *et al.*, 2005). In this study fluorescein isothiocyanate-labelled albumin (FITC-albumin) was detected in the columnar cells of the midgut and not the paracellular space, suggesting that paracellular transport was not occurring. In addition to this observation it was demonstrated that lumen-to-haemolymph flux was greater than in the reverse direction, which is unexpected for paracellular transport where the flux would be expected to be equivalent in both directions and suggests some form of active transport is occurring preferentially into the haemolymph from the gut lumen.

Further evidence that transepithelial transport in the *B. mori* midgut is energy-dependent was demonstrated by inhibition of transport when the midgut was incubated at 4 °C. Subsequently Casartelli *et al.* suggest that transepithelial transport is carried out by a transcytosis-like mechanism. A paracellular transport pathway would also be expected to exhibit a linearly

proportional relationship to the concentration in the luminal space. This was not observed by the work of Casartelli *et al.*

It seems likely that the transepithelial transport of proteins across the midgut is highly variable. There are vast differences in the structures of midgut lumens across different insect species, which almost certainly will facilitate different mechanisms of transport. Similarly, there is likely to be great variation in the movement of different types of proteins across the midgut, dependent on factors such as rates of degradation in the digestive system and the strength of interactions with the gut surface proteins or, in the case of plant lectins, glycoproteins. Jeffers *et al.* demonstrated that in the case of the tobacco budworm, *Heliothis virescens*, the uptake of anti-BSA into the haemolymph was approximately double that of BSA after 16 hours feeding (Jeffers, L. A., *et al.*, 2005). This was despite the molar concentration of anti-BSA (150 kDa) being less than half that of BSA (66 kDa) in the diets tested. This shows clearly that accumulation of proteins in the haemolymph is not simply dependent upon molecular weight or concentration of protein. In addition it is pointed out that as BSA and anti-BSA are not components of budworm diet, the mechanism of transport is not receptor specific.

2 Materials and Methods

2.1 Materials

All chemicals and reagents were supplied by Sigma Chemical Company (St. Louis, USA) unless stated. Chemicals and reagents were of analytical grade, or best commercially available. A head of *Allium sativum* bulbs was purchased from a local shop.

2.1.1 Commonly Used Media

LB broth	1% (w/v) NaCl, 1% (w/v) tryptone, 0.5% (w/v) yeast extract prepared with distilled water.
LB agar	LB broth plus 2% (w/v) Bacto-agar (Difco, http://www.bd.com/ds/).
LSLB broth	0.5% (w/v) NaCl, 1% (w/v) tryptone, 0.5% (w/v) yeast extract prepared with distilled water.
LSLB agar	LSLB broth plus 2% (w/v) Bacto-agar.
SOC medium	2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 20 mM glucose prepared with distilled water pH 7.0.

2.1.1.1 Yeast Culture Media

YPG media	1% (w/v) yeast extract, 2% (w/v) tryptone, 4% (v/v) glycerol prepared with distilled water.
YPG agar	YPG media plus 2% (w/v) Bacto-agar.

2.1.2 Commonly Used Buffers

2.1.2.1 DNA Electrophoresis

TAE (50x)	2 M Tris/Acetic Acid pH 7.7, 50 mM EDTA
-----------	---

DNA gel loading buffer	10 mM Tris/HCl pH 8.0, 10 mM EDTA, 30% (w/v) glycerol, 0.1% (v/v) Fast Orange G, prepared with distilled water.
------------------------	---

2.1.2.2 SDS-PAGE Buffers, Solutions and Size Markers

2x SDS sample buffer	0.2 M Tris/HCl (pH 6.8), 20% (v/v) glycerol, 2% (w/v) SDS, 0.002% (w/v) bromophenol blue.
----------------------	---

5x SDS sample buffer	0.5 M Tris/HCl (pH 6.8), 50% (v/v) glycerol, 5% (w/v) SDS, 0.005% (w/v) bromophenol blue.
----------------------	---

Acrylamide	30 g acrylamide, 0.8 g bisacrylamide/100ml.
------------	---

Resolving buffer	3.0 M Tris/HCl pH 8.8.
------------------	------------------------

Stacking buffer	0.5 M Tris/HCl pH 6.8.
-----------------	------------------------

Reservoir buffer (10x)	0.25 M Tris/HCl pH 8.3, 1.92 M glycine, 1% (w/v) SDS.
------------------------	---

Stain	40% (v/v) MeOH, 7% (v/v) glacial acetic acid, 0.05% (w/v) Kenacid blue.
-------	---

Destain	40% (v/v) MeOH, 7% (v/v) glacial acetic acid.
---------	---

Silver staining reagents:

Fixative	40% (v/v) EtOH, 10% (v/v) glacial acetic acid, 50% distilled water.
----------	---

Wash solution	30% EtOH, 70% distilled water.
---------------	--------------------------------

Thiosulphate reagent	0.02% (w/v) sodium thiosulphate in distilled water.
----------------------	---

Silver nitrate reagent	0.2% (w/v) silver nitrate, 0.02% (v/v) formaldehyde (37% solution) in distilled water.
------------------------	--

Developer	3% (w/v) sodium carbonate, 0.05% (v/v) formaldehyde (37% solution), 0.0005% (w/v) sodium thiosulphate in distilled water.
-----------	---

Stop reagent	0.5% (w/v) glycine in distilled water.
--------------	--

Protein Deglycosylation Buffer:

5x Assay buffer	250 mM NaHPO ₄ , pH 7.5.
Denaturation buffer	2% (w/v) SDS, 1 M βMercaptoethanol.

Protein Size Markers:

SDS 7 (Sigma)	66 kDa Bovine albumin
	45 kDa Egg albumin
	36 kDa Glyceraldehyde-3-phosphate
	29 kDa Carbonic anhydrase bovine erythrocytes
	24 kDa PMSF-treated trypsinogen
	20 kDa Soybean trypsin inhibitor
	14 kDa α-lactalbumin
Mark 12 (Invitrogen)	200 kDa Myosin
	116.3 kDa β-galactosidase
	97.4 kDa Phosphorylase B
	66 kDa Bovine albumin
	55.4 kDa Glutamic dehydrogenase
	36.5 kDa Lactate dehydrogenase
	31.0 kDa Carbonic anhydrase
	21.5 kDa Trypsin inhibitor
	14.4 kDa Lysozyme
	6.0 kDa Aprotinin
	3.5 kDa Insulin B chain
	2.5 kDa Insulin A chain

2.1.2.3 Western Blotting

Towbin's buffer	25 mM Tris
	20% (v/v) MeOH
	192 mM glycine

Ponceau S stain	0.1% Ponceau S
	5% Acetic Acid

2.1.2.4 Immunoblotting

Blocking solution	5% Non-fat milk powder, 1 x PBS, 1% Tween-20.
Anti-Sera solution	5% Non-fat milk powder, 1 x PBS, 0.1% Tween-20.
BSA blocking solution:	5% BSA, 1 x PBS, 1% Tween-20.
PBS (10x)	0.015 M KH_2PO_4 , 0.08 M Na_2HPO_4 , 1.37 M NaCl.
PBS-T	1 x PBS, 0.1% Tween-20.

2.1.2.5 ELISA Buffers

All buffers made up to 1 L in dH_2O .

COATING BUFFER:

0.1 M bicarbonate buffer pH 9.6

Na_2CO_3	4.24 g/l
NaHCO_3	5.04 g/l

WASH BUFFER:

10mM phosphate buffer pH 7.5

NaCl	8.766 g/l
$\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$	0.415 g/l
$\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$	0.148 g/l
Tween-20	1 ml

BLOCKING BUFFER:

Wash buffer + 2% Milk Powder

OPD SUBSTRATE BUFFERS:

A. 0.1 M citric acid 21.014 g/l

B. 0.2 M Na₂HPO₄·2H₂O 35.598 g/l

OPD SUBSTATE SOLUTION:

12.15 ml A + 12.85 ml B.

Add 25 mg OPD (O-phenylenediamine; Sigma), pH to 5.0

Then add 20 µl 30% H₂O₂

STOP BUFFER:

0.5 M H₂SO₄

1.33 ml conc. H₂SO₄ per 50 ml dH₂O

2.2 Standard Molecular Biological Techniques

All standard techniques carried out are common practice in the Durham University School of Biological and Biomedical Sciences and were based upon protocols from Molecular Cloning: A Laboratory Manual (Sambrook, J., Russell, D. W., 2001). Commercial kits were used according to the manufacturer's instructions unless otherwise stated.

2.2.1 Oligonucleotide Synthesis

Oligonucleotide primers were synthesized by either Sigma or TAGN (<http://www.tagn.co.uk/>).

2.2.2 Genomic DNA Extraction

Genomic DNA was extracted from *Allium sativum* bulbs using a Genelute Plant Genomic DNA extraction kit. Genomic DNA quality was assessed by 0.8% agarose gel electrophoresis.

2.2.3 Agarose Gel Electrophoresis

Separation of DNA by electrophoresis in agarose gel is described in Molecular Cloning (Sambrook, J., Russell, D. W., 2001). DNA was electrophoresed in gels containing agarose, typically 1% (w/v) in TAE buffer containing 0.5 µg/ml EtBr to allow visualization under UV light. DNA samples were made up to contain 1 x DNA gel loading buffer before loading. Gels were run at 50 – 100 V in TAE buffer containing 0.5 µg/ml EtBr. *Eco47I*-digested λDNA was used as a molecular size marker. Gels were photographed under UV light on a Gene Flash (Syngene www.syngene.com) Bio Imager.

2.2.4 Purification of DNA from Agarose Gel

DNA was purified from agarose gel using a Perfectprep Gel Cleanup kit (Eppendorf www.eppendorf.com) according to the manufacturer's instructions. A further dry spin was included to aid removal of residual ethanol where fragments were to be used in downstream subcloning procedures.

2.2.5 Bacterial Culture

For routine small-scale molecular biological operations liquid bacterial cultures were grown in 10 ml Luria-Bertani (LB) using single colonies picked from LB agar plates. Liquid cultures were grown at 37 °C overnight (16 hours) on a rotary shaker (approx. 200 rpm). Where appropriate, antibiotics for selection were added to culture media. In the case of bacteriological agar, antibiotics were added when the melted agar was at an approximate temperature of 50 °C. Where Zeocin was used as an antibiotic (Invitrogen www.invitrogen.com) LB media was replaced by LSLB.

2.2.6 Competent Cells

Competent cells of *Escherichia coli* used for plasmid transformation were chemically competent (Invitrogen) TOP10 cells.

2.2.7 Transformation of *E. coli*

Transformation of DNA into *E. coli* was performed following standard procedures (Sambrook, J., Russell, D. W., 2001). For chemically competent cells, a tube containing

50 µl of cells was removed from -80 °C storage and thawed on ice. 2 µl of DNA was added to the cells and gently mixed. After incubation on ice for 5-30 minutes cells were heat-shocked for 30 seconds at 42 °C. After heat shock, 250 µl of SOC medium was added and the tubes were incubated at 37 °C for 1 hour. The resulting suspension was plated out on LB-agar containing appropriate selection on two plates (50 µl and 250 µl respectively). When plating out plasmids containing the β galactosidase gene, 40 µl of X-Gal was spread on the agar plate and allowed to dry prior to inoculation with bacterial suspension to allow for blue-white screening. Plates were incubated at 37 °C for 16 hours.

2.2.8 Isolation of Plasmid DNA

Plasmid isolation was carried out based upon the alkaline lysis procedure. The Wizard *Plus* SV Miniprep DNA Purification System (Promega www.promega.com) was used for DNA minipreps according to supplied protocol.

2.2.9 Restriction Endonuclease Digestion of DNA

Restriction enzyme digestion was carried out using commercially available restriction endonuclease enzymes from a number of suppliers (Promega, New England Biolabs www.neb.com, Roche www.roche-applied-science.com, Fermentas www.fermentas.com). Typically, analytical digests were carried out in a total volume of 20 µl containing 5 µl of plasmid DNA. Digests for subcloning applications were typically carried out in a 50 µl volume containing 25 µl plasmid DNA. Digestion reactions were incubated for 1-3 hours at 37 °C. Where total digestion of DNA was required, 1 µl aliquots of the reaction mixture were taken every hour until complete digestion could be visualised after agarose gel electrophoresis.

2.2.10 Ethanol Precipitation of DNA

0.1 volumes of 3 M NaAc and 1 µl glycogen were added to solutions containing DNA. 2 volumes of ice-cold EtOH were added and the resulting mixture was vortexed thoroughly. Samples were incubated overnight at -20 °C and then centrifuged at 12000 rpm for 30 minutes at 4 °C. The supernatant was removed and 200 µl of 70% EtOH

added to the sample followed by centrifugation as previously described for 5 minutes. The supernatant was removed and the pellet briefly placed in a dessicator to dry. The pellet was then resuspended in distilled water.

2.2.11 DNA Ligation

DNA fragments were ligated in a 10 μ l reaction volume containing commercially available T4 DNA ligase and appropriate buffer (Promega). Digested and purified vector and insert DNA were added in approximate 1:3 and 1:1 ratios and the reaction made up to 10 μ l with distilled water. Ligation reactions were incubated at 16 °C overnight prior to transformation into *E. coli*.

2.2.12 DNA Amplification using Polymerase Chain Reaction (PCR)

PCR was performed using standard conditions. Reactions (typically 50 μ l) contained reaction buffer containing 1.5 mM MgCl₂, 0.2 M each of dATP, dCTP, dGTP and dTTP, 1 μ l each of 5' and 3' oligonucleotide primer, DNA template (50-100 ng) and 1 unit of Expand High Fidelity Polymerase (Roche, www.roche.com) or Phusion polymerase (www.finnzymes.fi) for cloning of expression constructs or *Taq* polymerase (for PCR-screening and colony PCR). Typically, for amplification involving plasmid templates, the PCR reaction consisted of an initial denaturation step of 10 minutes at 94 °C, 25 cycles of denaturation at 94 °C for 30 seconds, annealing at 50-60 °C for 45 seconds, extension at 72 °C for 45-60 seconds and a final extension step of 7 minutes at 72 °C. Reactions were carried out in an Applied Biosystems GeneAmp PCR system 2400 (www.appliedbiosystems.com), or a Thermo Hybaid PCR Sprint Temperature Cycling System (www.thermohybid.com).

2.2.13 Ligation of PCR Products into Cloning Vector

For storage and manipulation of PCR products, DNA from correctly sized bands was extracted from agarose gel after analysis and ligated into TOPO TA cloning vector (Invitrogen, www.invitrogen.com) transformed into *E. coli* TOP10 cells according to the manufacturer's instructions.

2.2.14 Colony PCR

Colony PCR was used as a method for rapid screening of recombinant plasmids. Ten to twenty colony transformants from recombinant *E. coli* or *P. pastoris* were picked and resuspended in 10 µl distilled water, this was then used in standard PCR reaction as previously described.

2.2.15 DNA Sequencing

DNA sequencing was carried out by the DNA Sequencing Laboratory, Durham University, using an Applied Biosystems 377 DNA Sequencer XL using standard double-stranded DNA templates and dye-dideoxy termination. Sequence data was obtained using standard vector primers or custom designed sequencing primers created for verification of correct sequences in expression constructs.

2.2.16 Glycerol Stocks

Single colonies of recombinant *E. coli* or *P. pastoris* containing DNA plasmids were inoculated into LB broth or YPG respectively containing appropriate antibiotic selection and grown overnight at 37 °C for *E. coli* or 30 °C for *P. pastoris*. 500 µl of liquid culture was added to 1.5 ml of sterile 60% glycerol (v/v) in 2 ml glass vials (BDH, www.bdh.com), vortexed and stored at -80°C.

2.3 Standard Biochemical Techniques

2.3.1 SDS-PAGE

Protein samples were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), using 12.5%, 15% or 17.5% resolving gel (12.5%, 15% or 17.5% (w/v) Protogel (37.5:1 acrylamide:bisacrylamide) (National Diagnostics, www.nationaldiagnostics.com), 0.375 M Tris/HCl (pH 8.8), 0.1% (w/v) SDS, 0.075% (w/v) ammonium persulphate, 0.05% (v/v) TEMED (*N, N, N', N'*-tetramethylethylenediamine) and 2.5% stacking gel (2.5% Protogel, 0.125 M Tris/HCl

(pH 6.8), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate and 0.075% (v/v) TEMED according to Laemmli (Laemmli, U. K., 1970). Minigels (8 x 10cm) were prepared and run in 1 x reservoir buffer at 70 V through the stacking gel and 120 V through the resolving gel using an ATTO AE-6450 gel tank apparatus (Genetic Research Instrumentation Ltd. <http://www.gri.co.uk/>). Protein samples were prepared by adding SDS sample buffer to a 1 x concentration (from a 2 x or 5 x stock solution depending on the volume of protein solution in the sample), treated with 10% β -mercaptoethanol and heated in boiling water for 5 minutes before loading onto gel. A molecular weight marker was used to calibrate gels (either SDS-7 or Mark 12, Invitrogen). Proteins were visualised on gel using either Kenacid blue stain solution or silver staining.

2.3.2 Silver Staining

Silver staining of SDS-PAGE gels was carried out to a protocol modified from the standard protocol described by Blum (Blum, H. *et al.*, 1987). The gel underwent 1 hour incubation in 50 ml fixative, two 20 minute washes in 50 ml wash solution followed by a single 20 minute wash in 50 ml distilled water. This was followed by 1 minute incubation in 50 ml sensitizer reagent followed by three 20 second washes in 50 ml distilled water, 20 minute incubation in 50 ml silver nitrate solution and three 20 second washes in 50 ml distilled water. The gel was then incubated in 50 ml developer solution until proteins were resolved sufficiently and background had not begun to appear on the gel, the gel was then washed in approximately 200 ml distilled water before being incubated in 50 ml stop reagent for 5 minutes.

2.3.3 SYPRO Ruby Staining

SYPRO Ruby staining was carried out on SDS-PAGE gels following the basic protocol outlined in the manufacturer's instructions (Invitrogen www.invitrogen.com). Briefly, a fixing step was carried for 30 minutes in 100 ml fixing solution (50% MeOH, 7% acetic acid). This step was then repeated before an overnight incubation in 60 ml SYPRO Ruby staining solution. The gel was then washed for 30 minutes in wash solution (10% methanol, 7% acetic acid), followed by 2 x 5 minute washes in dH₂O. The gel was then ready to be imaged using a Typhoon Imager (Amersham biosciences).

2.3.4 Immunoblotting

Proteins were transferred from polyacrylamide gel to nitrocellulose membrane (Protran BA85, Whatman Ltd. www.whatman.com) by semi-dry blotting. Six sheets of 3MM paper (Whatman Ltd. www.whatman.com) and a piece of nitrocellulose membrane were cut to the same size as the gel and soaked for 10 minutes in Towbins buffer. The blotting sandwich was then prepared in an ATTO blotting apparatus (Genetic Research Instrumentation Ltd. <http://www.gri.co.uk/>) as follows: ANODE; three sheets of 3MM paper; nitrocellulose membrane; gel; three sheets of 3MM paper; CATHODE. Electroblotting was carried out at 150-200 mA for 1 hour. For visualization of protein standard marker transfer prior to immunodetection, the nitrocellulose membrane was soaked in Ponceau S stain for 5 minutes and rinsed thoroughly in water.

For immunodetection, the nitrocellulose membrane was blocked for either 1 hour at room temperature or at 4 °C overnight in blocking buffer. Following 3 x 5 minute washes in antisera buffer, primary antibody was diluted 1:3000 (anti-myc, anti-ASA), or 1:10000 (anti-avidin) in 10 ml anti-sera buffer and the membrane incubated overnight at 4 °C. The membrane was then again washed in anti-sera buffer for 3 x 5 minutes at room temperature. Goat anti-rabbit (for anti-avidin and anti-ASA), or goat anti-mouse (anti-myc) IgG horseradish peroxidase conjugate (Biorad, www.biorad.com) was used as a secondary antibody in 10 ml blocking buffer at a 1:5000 dilution and incubated with the membrane at room temperature for 3 hours. This was followed by 1 x 5 wash in anti-sera buffer then 1 x 15 minute and 2 x 5 minute washes in PBS-T and several rinses in distilled water.

Enhanced chemoluminescence (ECL) reagents (GE healthcare, www.gehealthcare.com) were used for detection according to the manufacturer's instructions and proteins were visualized by exposure to X-ray film (Fuji SuperRX, Fuji Photo. Film Ltd, www.fujifilm.co.uk). Autoradiographs were either developed using an automatic developer (Xograph Imaging Systems Compact X4, www.xograph.com) or manually.

2.3.5 Protein Concentration Determination

The protein content of solutions was determined using the BCA Protein Assay Kit (Pierce, www.piercenet.com) according to the manufacturer's instructions for the microplate protocol.

2.3.6 Deglycosylation of Protein Samples

50 µl deglycosylation reactions were carried out as follows. 50-100 µg of protein was added to 35 µl distilled water. 10 µl of 5x assay buffer and 2.5 µl denaturation buffer were added and boiled for 5 minutes. Once cooled 2.5 µl Triton X-100 was added and mixed. 2 µl of N-glycosidase F from *Chryseobacterium meningosepticum* (Calbiochem, www.emdbiosciences.com) was added and the reaction was incubated overnight at 4 °C.

2.3.7 Dialysis of Protein Samples

Pooled fractions from chromatographic peaks were dialysed against two changes of 50 mM ammonium hydrogen carbonate overnight at 4 °C. Dialysis tubing was prepared by 10 minutes boiling in a solution of 2% (w/v) sodium bicarbonate and 1 mM EDTA (pH 8.0).

2.3.8 Lyophilization of Protein Samples

Following dialysis protein solutions were snap-frozen in liquid nitrogen and placed on a vacuum freeze-dryer overnight until lyophilization occurred.

2.4 Production of Recombinant Proteins in *P. pastoris*

2.4.1 Transformation of *P. pastoris*

Transformation of *P. pastoris* X33 and SMD1168 strains was carried out by isolating recombinant pGAPzαA plasmid DNA from 60 ml of overnight bacterial culture. Plasmid DNA was eluted into 200 µl distilled water and 190 µl was digested with *BlnI* restriction enzyme (Roche, www.roche.com). The digested DNA was ethanol precipitated overnight

and resuspended in 11 µl distilled water. A 1 µl aliquot was taken and run on agarose gel alongside the remaining 1 µl of undigested plasmid DNA with 5 µl Eco471/λ DNA marker to ensure the DNA had digested correctly. The remaining 10 µl of *BlnI* digested plasmid DNA was used in the transformation. Transformation of both X33 and SMD1168 were carried out using an Easycomp *Pichia pastoris* Kit (Invitrogen, www.invitrogen.com). The protocol used differed slightly from that of the manufacturer's instructions in that 10 µl of DNA was used where the manufacturer recommended 5 µl. Subsequently, all volumes in the manufacturer's instructions were doubled apart from the final resuspension of cells, which was made in 150 µl. Putatively transformed cells were plated out on YPD agar containing 100 µg/ml Zeocin and incubated for 3-5 days at 30 °C. Colonies were then transferred to fresh selection media and screened for recombination plasmid by colony PCR.

2.4.2 Expression Screening of Transformed *P. pastoris* Colonies

To determine which transformant colonies were expressing recombinant lectins in the greatest abundance, 10 ml YPD 100 µg/ml Zeocin cultures of all positive colonies were grown for three days. To precipitate proteins 0.6% (w/v) ammonium sulphate was added to each culture and incubated overnight at 4 °C on a roller shaker. The cultures were then spun for 15 minutes at 20000 g. The pellet was resuspended in 100-300 µl and a 20 µl aliquot was analysed by western blotting.

2.4.3 Overexpression of Recombinant Proteins in *P. pastoris*

For overexpression of ASAI and ASAI in *P. pastoris* cells were grown in a 5-litre benchtop fermenter (New Brunswick Scientific, <http://www.nbsc.com/>). Initially, three 100 ml starter cultures were inoculated in YPG media containing 100 µg/ml Zeocin. Starter cultures were incubated at 30 °C for 3 days with shaking at 200 rpm. Starter cultures were then used to inoculate 3.75 litres of Basal Salt Medium (Higgins, D. R., Creggs, J. M., 1998). Fermentation was carried out at 30 °C with a constant dissolved oxygen level of 30% at pH 5.0. When cell density became such that 30% dissolved

oxygen could not be sustained by agitation alone, a constant feed of 50% glycerol (v/v with distilled water) was initiated (~1 ml/minute). When the glycerol feed was exhausted the run was continued until the dissolved oxygen level began to rise, indicating growth had stopped and the run was terminated.

2.4.4 Recovery of Protein extract from *P. pastoris*

Total secreted protein was recovered from *P. pastoris* cell suspension by 30 minutes centrifugation at 4 °C at a relative centrifugal force of 9000 g. Recombinant protein was collected in the supernatant due to the α -factor secretory sequence of pGAPz α A. For the purposes of later purification, approximate salt concentration of the supernatant was estimated to be 0.3 M.

2.5 Purification of Recombinant Garlic Lectins from *P. pastoris* Culture Supernatant

2.5.1 Purification of Recombinant Lectins by Hydrophobic-Interaction Chromatography

Crude culture supernatant was taken and NaCl added to a final concentration of 2 M and filtered through a 0.7 μ m filter (GF-F, Whatman, www.whatman.com) to prepare the sample for loading. A phenyl-sepharose matrix (Amersham Biosciences, www.amersham.com) column was prepared and equilibrated with 2 M NaCl. The sample loaded at 4 °C and eluted across a 2 M to 0 M NaCl gradient at a flow rate of 2 ml/min collecting 5 ml fractions. Typically the gradient was carried out over 120 ml with a major peak eluting at the 0 M NaCl stage.

2.5.2 Purification of Recombinant Lectins by Anion-Exchange Chromatography

For purification of recombinant lectins by anion-exchange chromatography crude culture supernatant was made up into a final concentration of 20 mM diaminopropane (DAP), pH 9.0, and filtered through a 0.7 μ m filter. A Q-sepharose matrix column (Amersham) was

prepared and equilibrated with 20 mM DAP containing 0 M NaCl. The protein sample was then loaded at 4 °C and eluted across a 0 M to 0.5 M NaCl gradient at a flow rate of 2 ml/min collecting 5 ml fractions. Typically the gradient was carried out across 120 ml with a major peak eluting at 0.2 M to 0.3 M NaCl.

2.5.3 Purification of Recombinant Lectins by Cation-Exchange Chromatography

For purification by cation-exchange chromatography the NaCl present in crude culture supernatant was diluted out after it was observed that without such a reduction in NaCl concentration in the load sample, binding to the matrix was not occurring. Subsequently a 10-fold dilution of crude culture supernatant into 50 mM sodium acetate (NaAc), pH 4.0, was carried out. An S-Sepharose matrix (Amersham) column was prepared and equilibrated with 50 mM NaAc, pH 4.0. The protein sample was loaded at 4 °C and eluted over a gradient of 0 M to 0.5 M NaCl at a flow rate of 5 ml/min collecting 12.5 ml fractions. Typically the gradient was carried out over 300 ml with a major peak eluting at around 0.2 M NaCl.

2.5.4 Purification of Recombinant Lectins by Metal-Affinity Chromatography

Recombinant lectins were purified using nickel-affinity chromatography on a 1 ml HiTrap Ni²⁺ affinity column (Amersham). Crude culture supernatant was made up to a final concentration of 50 mM sodium phosphate (NaP), pH 7.0, and passed through a 0.7 µm filter. The column was equilibrated in 50 mM sodium phosphate and the protein sample loaded. After loading the column was washed with 50 mM NaP containing 20 mM imidazole to remove non-specifically bound protein. Bound protein was eluted in 50 mM NaP containing 200 mM imidazole. The elution fraction was immediately diluted and dialysed to prevent precipitation. For purification of proteins that precipitated at pH 7.0, a loading buffer of 100 mM NaAc, pH 4.0 was used (washing and elution buffers also contained 100 mM NaAc and imidazole concentrations as given above). The flow rate used was 1 ml/min for all affinity purifications as recommended in the manufacturer's instructions.

2.6 Haemagglutination Assays of Recombinant Proteins

Activity of recombinant lectins was measured by haemagglutination assay. Rabbit blood supplied in heparin was diluted 25-fold with ice cold PBS (pH 7.4) and spun for 1 minute at 300 g to wash the erythrocytes. The pellet was then resuspended in 1ml ice-cold PBS at an approximate final concentration of 4% erythrocytes. Assays were carried out in round-bottomed or v-shaped 96-well microtitre plates, where each test well received 50 μ l PBS. 100 μ l of 0.1mg/ml GNA stock was serially diluted as a positive control. 50 μ l of PBS was added to negative control wells. Recombinant lectin samples were made up to 2.5 mg/ml and serially diluted as for positive controls. Agglutination assays were carried out for 3 hours at room temperature or overnight at 4 °C. Agglutination activity was visualized as a cross-linking of erythrocytes within the well, where agglutination had not occurred erythrocytes formed a pellet.

2.7 Insect Culture

2.7.1 *N. lugens* stock culture

A longstanding laboratory culture of rice brown plant hopper (*Nilaparvata lugens*) were reared on 20-day rice plants (var. TN1) subject to a constant temperature of 28 °C, 90% relative humidity and 16-hour light, and 8-hour dark regimen.

2.7.2 *N. lugens* artificial diet

N. lugens artificial diet MMD-1 was formulated according to Mitsuhashi, 1974. This formulation has been demonstrated to be the most appropriate for use in *N. lugens* artificial diet feeding trials (Powell, K. S., 1993). The formulation of this diet is shown in **Error! Reference source not found..** The amino acids cysteine and tyrosine required dissolving in a minimal volume of 1 M HCl and riboflavin was heated gently in dH₂O prior to addition to the diet mixture. Diet was filter-sterilized using a 0.2 μ M pore filter and subsequent manipulations were carried in a sterile laminar flow hood.

2.7.3 *N. lugens* Artificial Diet Bioassay

Third instar insects were taken from rice plants by tapping the stem and collected in petri dishes. Insects were transferred into feeding chambers using a fine paintbrush; ten insects were placed in each feeding chamber and two replicates were set up for each treatment. Diet sachets were prepared by stretching a layer of Parafilm M to approximately four times its starting size and placing it over a small petri dish (diameter 30 mm), 200 μ l of artificial diet was dispensed onto the parafilm and a second layer of similarly stretched parafilm was placed over the diet, creating a sachet. Feeding chambers were placed in a growth room optimized for *N. lugens* survival (see section 2.7.1). Artificial diet was replaced every other day to prevent bacterial contamination. Numbers of surviving nymphs were recorded at the same time every day.

Ingredient	mg/L	Ingredient	mg/L
K ₂ HPO ₄	7500	Threonine	1400
MgSO ₄ •7H ₂ O	1230	Tryptophan	800
Sucrose	50000	Tyrosine*	400
Alanine	1000	Valine	800
Arginine•HCl	2700	Thiamine•HCl	25
Asparagine	5500	Riboflavin [#]	5
Aspartic Acid	1400	Nicotinic Acid	100
Cysteine*	400	Pyroxidine•HCl	25
Glutamic Acid	1400	Folic Acid	5
Glutamine	1500	Calcium pantothenate	50
Glycine	800	Myo-inositol	500
Histidine	800	Choline chloride	500
Isoleucine	800	Biotin	1
Leucine	800	Sodium ascorbate	1000
Lysine•HCl	1200	FeCl ₃ •6H ₂ O	22.28
Methionine	800	CuCl ₂ •4H ₂ O	2.68
Phenylalanine	400	MnCl ₂ •6H ₂ O	7.93
Proline	800	ZnCl ₂	11.88
Serine	800	CaCl ₂ •2H ₂ O	31.15

Table 2.1 Composition of *N. lugens* artificial diet MMD-1. Cysteine and tyrosine were dissolved in a small volume of 1 M HCl prior to addition to the diet mixture. Riboflavin was dissolved by gentle heating in dH₂O.

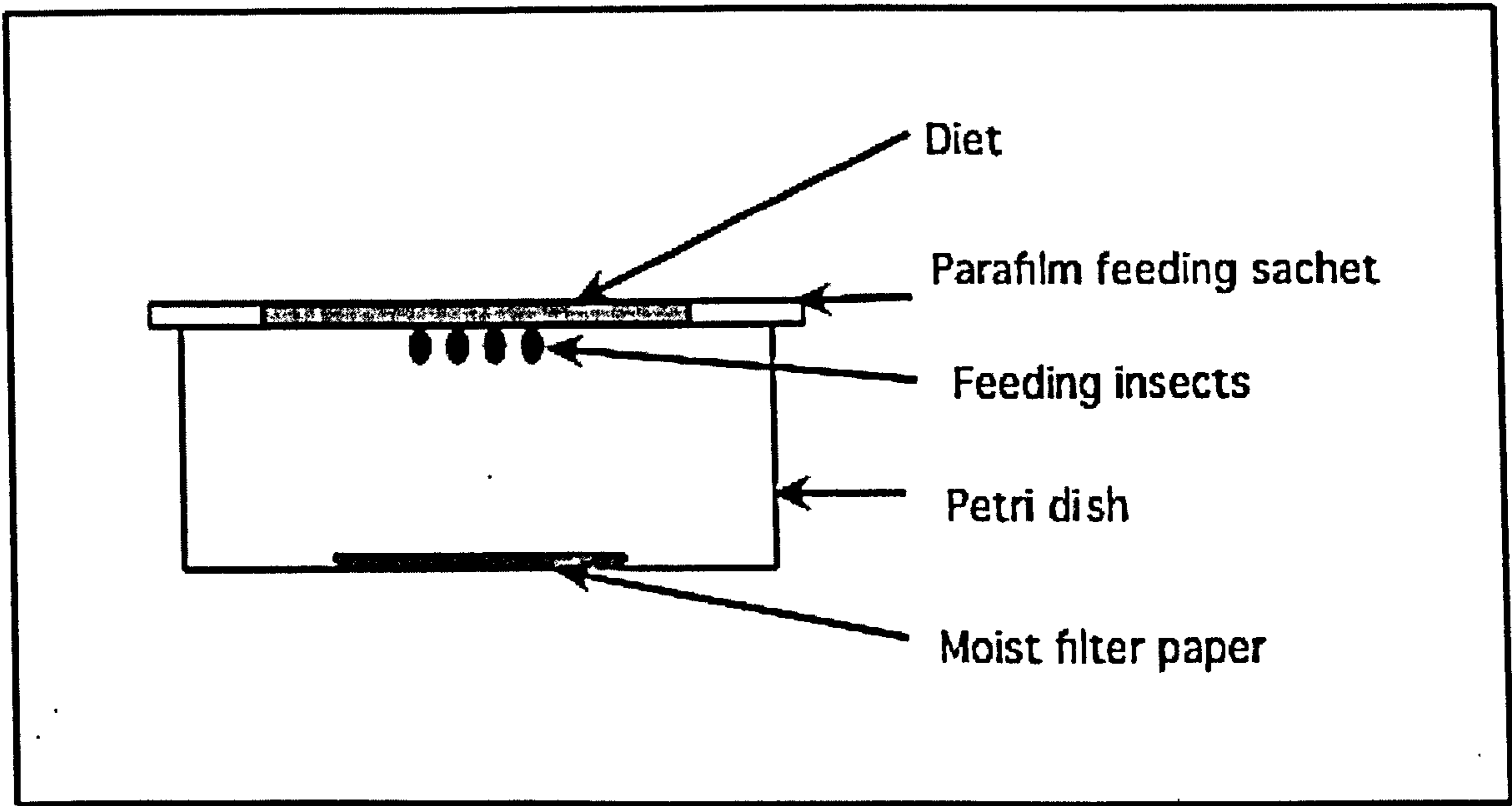


Figure 2.1 Feeding chamber used for *N. lugens* artificial diet bioassays. Humidity is maintained by moist filter paper in the chamber. 200 µl of artificial diet was added to each chamber.

2.7.4 *A. pisum* stock culture

The pea aphid *Acyrtosiphon pisum* clone LL01 was acquired from Prof. Angela Douglas (University of York) and maintained on pre-flowering *Vicia faba* cv. The Sutton at 16 °C with a 16 hour light, 8 hour dark regimen.

2.7.5 *A. pisum* artificial diet

Standard aphid diet was prepared (Prosser, W. A. and Douglas, A. E. 1992) using the stock solutions described below:

The following stocks were prepared in advance and stored at -20 °C:

Mineral Stock (in 10 ml distilled water):

Compound	Milligrams in 10 ml
FeCl ₃ •6H ₂ O	11
CuCl ₂ •4H ₂ O	2
	4

MnCl ₂ •6H ₂ O	
ZnSO ₄	17

Aliquot into 0.1 ml lots and store.

Table 2.2 Stock solutions for the formulation of minerals for inclusion in standard aphid diet.

Vitamin Stock (Dissolve in 5 ml distilled water):

Compound	Milligrams in 5 ml
Biotin	0.1
D-Pantothenic Acid	5
Folic Acid	2
Nicotinic Acid	10
Pyroxidine	2.5
Thiamine	2.5
Choline	50
Myo-inositol	50

Aliquot into 0.5 ml lots and store.

Table 2.3 Stock solutions for the formulation vitamins for inclusion in standard aphid diet.

Amino Acid Stock (Dissolve in 50 ml distilled water):

Amino Acid	Milligrams to give 150 mM final concentration	Milligrams to give 2 x 150 mM final concentration
Alanine	50.8	101.6
Arginine•HCl	300.2	600.4
Asparagine•H ₂ O	213.9	427.8
Aspartic Acid	189.7	379.4
Cysteine•HCl	42.5	85.0

Glutamic Acid	123.6	247.2
Glutamine	241.1	482.2
Glycine	9.0	18.0
Histidine•HCl•H ₂ O	182.4	364.8
Isoleucine	114.1	228.2
Leucine	114.1	228.2
Lysine•HCl	158.9	317.8
Methionine	42.5	85.0
Phenylalanine	47.1	94.2
Proline	65.6	131.2
Serine	59.9	119.8
Threonine	103.6	207.2
Tryptophan	58.2	116.4
Tyrosine	10.9	21.8
Valine	101.9	203.8

Aliquot into 5 ml lots and store.

Table 2.4 Stock solutions for the formulation of amino acids for inclusion in standard aphid diet.

For 1 x standard aphid diet 0.1 ml of mineral stock solution, 0.5 ml of vitamin stock solution and 5 ml of amino acid stock solution were mixed and 1.7 g of sucrose (500mM final concentration), 10 mg ascorbic acid, 1mg citric acid and 20 mg MgSO₄•7H₂O added. 150 mg of K₂PO₄•3H₂O was dissolved 1 ml dH₂O and added to the diet mixture. pH was checked and verified to be 7.0-7.5 and then the diet mixture was made up to a total volume of 10 ml and filter-sterilized through a 0.2 µm filter. For diet containing antibiotic, 5 ml of diet was taken and 50 µl of a 5 mg/ml rifampicin stock solution was added prior to filter-sterilization. For 2 x standard aphid diet amino acid stocks were made at 2 x concentration. All other stock solutions and diet components were added in twice the amount for 1 x diet and the volume was adjusted to 10 ml accordingly.

2.7.6 *A. pisum* Artificial Diet Bioassay

Viviparous adult apterae were removed from pre-flowering *V. faba* using a fine paint brush and placed on a single cut leaf in a Blackman box for symbiotic trials or onto artificial diet containing rifampicin for aposymbiotic trials; after one day adult apterae were removed and day 0 nymphs were collected for inclusion in artificial diet feeding trials. Feeding chambers were set up similarly to those used for *N. lugens* trials but plastic rings of approximately 30 mm diameter and 20 mm depth, termed 'diet rings' were used in place of petri dishes and the diet sachet contained 100 µl of artificial diet.

2.7.7 Insect Dissection and Haemolymph Extraction

Aphid guts were dissected from adult insects in ice-cold 0.9% saline solution using fine forceps under a binocular microscope (Nikon, www.nikon.com). Haemolymph was extracted, typically from day 4-5 aphids, in mineral oil (Sigma) using pulled microcapillary tubes by removal of a single leg and collection of haemolymph from the wound site. Haemolymph was ejected from the microcapillary tubes using a P200 pipette (Gilson, www.gilson.com) with a cut pipette tip allowing the microcapillary tube to be inserted.

2.8 Aphid Gut Pull-Down Assays

2.8.1 Preparation of ASaII-Magnetic Beads

Amine-derivatized Magnabind magnetic beads (Pierce, www.piercenet.com) were cross-linked to purified ASaII produced in *P. pastoris* strain X33 using a Bis[sulfosuccinimidyl] suberate (BS³) cross-linker (Pierce). 0.5 ml of magnetic beads were washed 3 times in 0.5 ml PBS and the supernatant removed. After washing, 0.5 ml of an 8 mg/ml ASaII PBS solution was added to the beads. To cross-link 0.25 ml of a 5 mM BS³ PBS solution was added to make a final BS³ concentration of 1 mM. The reaction was incubated for 30 minutes at room temperature before 50 µl of 1 M Tris-HCl

(pH 7.5) stop solution was added. The beads were then washed 3 times with 0.5 ml PBS and stored at 4 °C. The amount of protein cross-linked to the beads was calculated from an initial spectrophotometric value at 280 nm and a value for the supernatant after the reaction had been stopped prior to the final wash steps. It was shown that 89% of the purified protein had been cross-linked to the beads, meaning the bead solution contained 7.2 mg/ml ASaII.

2.8.2 Pull-Down Assay Sample Preparation

A. pisum guts were dissected in 0.9% saline and stored in a microcentrifuge tube on ice. The gut tissue was then spun down and the saline solution removed. The pellet was then resuspended in 200 µl of ice-cold 50 mM HEPES buffer (pH 7.0) and homogenized with a mini-pestle. The homogenate was then spun down at 4 °C and 12000 g for 10 minutes. The supernatant was then removed and represented the soluble protein fraction of the *A. pisum* gut. The pellet was then washed once with 50mM HEPES and spun down as described. The pellet was then resuspended in 50mM HEPES + 0.1% Triton X-100 and homogenized. The homogenate was then incubated at 4 °C for 15 minutes with shaking. Finally the homogenate was spun down as described and the supernatant was removed and taken as the insoluble protein fraction of the *A. pisum* gut. All samples were stored at -20 °C until required.

2.8.3 Pull-Down Assays

The sample to be analysed was thawed and resuspended in 0.5 ml ice-cold PBS for soluble fractions or 0.5 ml ice-cold PBS containing 0.1% Triton X-100 for insoluble fractions and added to 50 µl of the magnetic bead solution described in section 2.7.1. This mixture was placed at 4 °C on a roller shaker overnight. Prior to analysis by SDS-PAGE the beads were collected by a magnetic stand and the initial supernatant removed and kept as the 'unbound' protein sample. 0.5 ml ice-cold PBS was then added and the beads incubated at 4 °C for 5 minutes before being collected and the supernatant removed, this process was repeated three times and the supernatants were taken as 'wash' fractions. After removal of the final wash fraction the beads were resuspended in 1 x

SDS-sample buffer and labeled the 'elution' fraction. Aliquots (typically 20 μ l) of the starting sample preparation, the unbound fraction, wash fractions and beads only were prepared for SDS-PAGE analysis alongside the elution fraction. Prior to loading on gel the elution fraction was spun down briefly in a bench-top centrifuge after boiling to prevent magnetic beads being loaded onto the gel.

2.9 MALDI-TOF Analysis

2.9.1 Analysis of Recombinant Lectins by MALDI-TOF

MALDI-TOF analysis was carried out on lyophilized ASAI and ASAII samples expressed in the X33 strain of *P. pastoris*. The procedure carried out for each was essentially as described (Cohen, S. L., Chait, B. T., 1997). Briefly, lyophilised samples were resuspended in matrix solution (α -cyano-4-hydroxy-cinnamic acid in 50% acetonitrile/0.05% trifluoroacetic acid (TFA)) and allowed to crystallise. 1 μ l of the solution was deposited onto the MALDI probe and left to air dry. MALDI-TOF was then performed using a Voyager-DE™ STR BioSpectrometry™ Workstation (Applied Biosystems, Warrington, UK). Peptide mass data was obtained and these masses were compared to the predicted amino acid sequences of ASAI and ASAII using the FindPept programme at www.expasy.org.

2.9.2 Peptide Mass Fingerprint (PMF) Analysis of ASAII Binding Proteins

Protein bands were extracted from an SDS-PAGE gel using a clean scalpel blade and removed to a well of a 96-well microtitre plate (Genomic Solutions, www.genomicsolutions.com). The microtitre plate was transferred to a ProGest Workstation (Genomic Solutions) and the sample was digested with trypsin using the ProGest long trypsin digestion protocol. Briefly, gel slices were equilibrated in 50 μ l of 50 mM ammonium bicarbonate and subsequently reductively alkylated with 10 mM DTT and 100 mM iodoacetamide, followed by destaining and desiccation with acetonitrile. Then the gel slices were rehydrated with 50 mM ammonium bicarbonate containing 6.6%

(w/v) trypsin (Promega) and digested overnight. Peptides were extracted using 50% (v/v) acetonitrile, 0.1% (v/v) TFA into a final volume of 50 μ l (2 x 25 μ l extractions). The resulting extracts are lyophilised and then resuspended in 10 μ l of 0.1% formic acid. The sample is applied to the MALDI probe using the 'thin film' method. 2 μ l of matrix solution (α -cyano-4-hydroxy-cinnamic acid in nitrocellulose/acetone), 1 μ l of sample is then added to this matrix solution film and allowed to air dry. The samples are washed *in situ* with 0.1% TFA and then left to dry again. MALDI-TOF PMF was then performed using a Voyager-DE™ STR BioSpectrometry™ Workstation (Applied Biosystems, Warrington, UK). De-isotoped & calibrated spectra were then used to generate peak lists, which were searched using MASCOT (www.matrixscience.com) mass spectrometry database search software and the NCBI non-redundant protein database.

Chapter 3

Production of Functional Recombinant ASAI & ASAI in *Pichia pastoris*

3.3 Introduction

The plant lectins are a heterogenous group of proteins. Attempts to understand the nature of these proteins has led to the expression of different lectins as recombinant proteins, a process which has met with varying degrees of success; the lectin and expression hosts chosen have led to differing levels of protein expression and recovery. In some cases recombinant lectins have activities comparable to biochemically isolated native proteins, whereas in others the recombinant protein has significantly lower activity, or is inactive or insoluble. Initially, bacterial expression systems were used routinely for recombinant protein expression, and expression of some lectins as recombinant proteins led to soluble proteins with comparable biological activity to native lectins. More recently eukaryotic yeast systems have been utilized to overcome instances where bacterial systems have failed to produce fully functional recombinant proteins. Eukaryotic expression systems, such as *Pichia pastoris*, are able to carry out many of the post-translational processing events which are not possible using bacterial systems.

Recombinant plant lectins have been expressed in bacterial systems since the mid-1980s (Stubbs, M. E., *et al.*, 1986) when a cDNA encoding lectin from garden pea (*Pisum sativum*) was produced in *Escherichia coli*. The recombinant lectin expressed by Stubbs *et al.* was not functional, and was only present in insoluble fractions of *E. coli* cell lysate. In order to recover active lectin, a lengthy procedure involving denaturation with guanidine hydrochloride, and renaturation by removal of the denaturant was necessary. Not all the protein refolded correctly and remained soluble, and not all the soluble protein was functionally active, since an affinity chromatography step based on lectin binding activity was necessary to remove inactive lectin. Since then, a number of other plant and fungal lectins have been expressed in *E. coli* and demonstrated to have agglutination activity (Chao, Q., *et al.*, 1994, Tateno, H., *et al.*, 2004). However, the recovery of many of these recombinant lectins has again been complicated by similar problems of insolubility at high expression levels. Recovery of recombinant lectins from insoluble cell

lysate by denaturation-renaturation protocols has produced active proteins from a number of plant species, and in some cases this activity has been demonstrably similar to that of the native lectin; for example *Pinellia ternata* agglutinin in native and recombinant forms has similar agglutination activities (Lin, J., *et al.*, 2003). Comparable agglutination activities were also observed between native and recombinant *Galanthus nivalis* agglutinin (GNA) (Longstaff, M., *et al.*, 1998). Some studies have found that recombinant plant lectins expressed in *E. coli* demonstrate lower agglutination activities than their native equivalent, notably the recombinant lectin from *Dolichos biflorus*, which showed decreased carbohydrate binding activity when compared to the native lectin (Chao, Q., *et al.*, 1994). The native lectin from *D. biflorus* is a glycoprotein, and the inability of *E. coli* to produce glycosylated proteins may contribute to the reduced activity of the recombinant lectin. Many plant lectins are glycoproteins with covalent N-linked carbohydrate groups attached to asparagine residues in the N-X-T/S peptide motif; N-glycosylation does not occur in plant lectins which lack this motif. The significance of glycosylation of plant lectins, which may affect the transport, stability and carbohydrate binding activity of these proteins, is not well understood, and may vary depending on the specific lectin and plant species being considered.

Other post-translational modifications of plant lectins besides glycosylation are common, and usually take the form of proteolytic cleavage of parts of the polypeptide chain. The presence and nature of post-translational proteolysis varies from species to species; in some cases C-terminal extensions are removed, whereas in others linker regions between polypeptides of the mature protein are excised. The most complex example of proteolytic modification so far identified is shown by the lectin Concanavalin A (ConA) isolated from *Canavalia ensiformis*. Pro-ConA is synthesized as a glycosylated precursor, which is in turn deglycosylated, cleaved on the carboxy side of asparagine residues and finally reformed as two peptide fragments (1-118 and 119-237) which are reversed and a peptide bond formed *de novo* (Bowles, D. J., *et al.*, 1986), (Sheldon, P. S., *et al.*, 1996). Despite this complex processing, active ConA has been produced in *E. coli* in its pro-ConA form, suggesting that post-translational modification is not essential for activity (Min, W., *et al.*, 1992). Similarly, pea lectin expressed in *E. coli* did not require the post-translational proteolysis observed *in planta* to be active (Stubbs *et al.*, 1986).

The methylotrophic yeast *Pichia pastoris* has been used successfully in the past to express plant lectins from *Galanthus nivalis* and *Phaseolus vulgaris* as soluble proteins (Raemaekers, R. J., *et al.*, 1999), (Baumgartner, P., *et al.*, 2002), (Baumgartner, P., *et al.*, 2003). The use of a eukaryotic expression system has overcome the problems of refolding and non-glycosylation of recombinant proteins expressed in bacterial systems. In order to aid downstream processing of recombinant proteins, an expression vector containing the secretory sequence from the *Saccharomyces cerevisiae* α -factor was used; constructs were assembled with the α -factor prepro-sequence fused in-frame to the N-terminal of the protein to be expressed, and the recombinant protein was subsequently secreted to the culture supernatant. The protein was soluble and fully functional. One potential drawback of the *P. pastoris* expression system that has been observed is endogenous protease activity in culture supernatant during the fermentation period, due to secreted aspartic proteases. Consequently, *pepA* genotype protease-deficient *P. pastoris* strains (such as SMD 1168) have been used as expression hosts, as an alternative to the wild-type strains (such as X-33).

The cDNAs encoding ASAI and ASAIL were first identified by van Damme *et al.* (van Damme, E. J. M., *et al.*, 1992) from total mRNA of garlic bulbs. A number of cDNAs for ASAI and ASAIL were identified and are found in the Genbank/EMBL Database (accession numbers M85174-M85177 and M85171-M85173 respectively). Native garlic bulb lectins have been purified using affinity chromatography on immobilized mannose, demonstrating the strongest affinity for α -(1,3)-linked mannosyl units. The relative activities of native ASAI and ASAIL in agglutination of rabbit erythrocytes have shown that ASAI has a 10-fold greater agglutination activity than ASAIL (van Damme, E. J. M., *et al.*, 1998). In contrast with the other lectins present in garlic, the garlic leaf lectin and the garlic root lectin, bulb lectins are highly abundant and developmentally regulated (van Damme, E. J. M., *et al.*, 1992) which suggests a role as storage proteins. Slight insecticidal effects have been observed using an ASAI/ASAIL mixture against *N. lugens* and *M. persicae* (Powell, K. S., *et al.*, 1995), (Sauvion, N., *et al.*, 1996).

In this chapter cloning and expression of recombinant garlic bulb lectins ASAI and ASAIL in *P. pastoris* is demonstrated and details of purification given. Recombinant lectins were expressed in wild-type and protease-deficient strains and analysis of the

overexpression products carried out by MALDI-TOF. The recombinant lectins produced were tested for *in vitro* biological activity and correct post-translational modifications. Consideration is given to the effects of C-terminal processing of ASAI and ASAIL. The exact nature of C-terminal processing *in planta* is unknown, but the importance of this modification on biological activity or oligomerization may be significant.

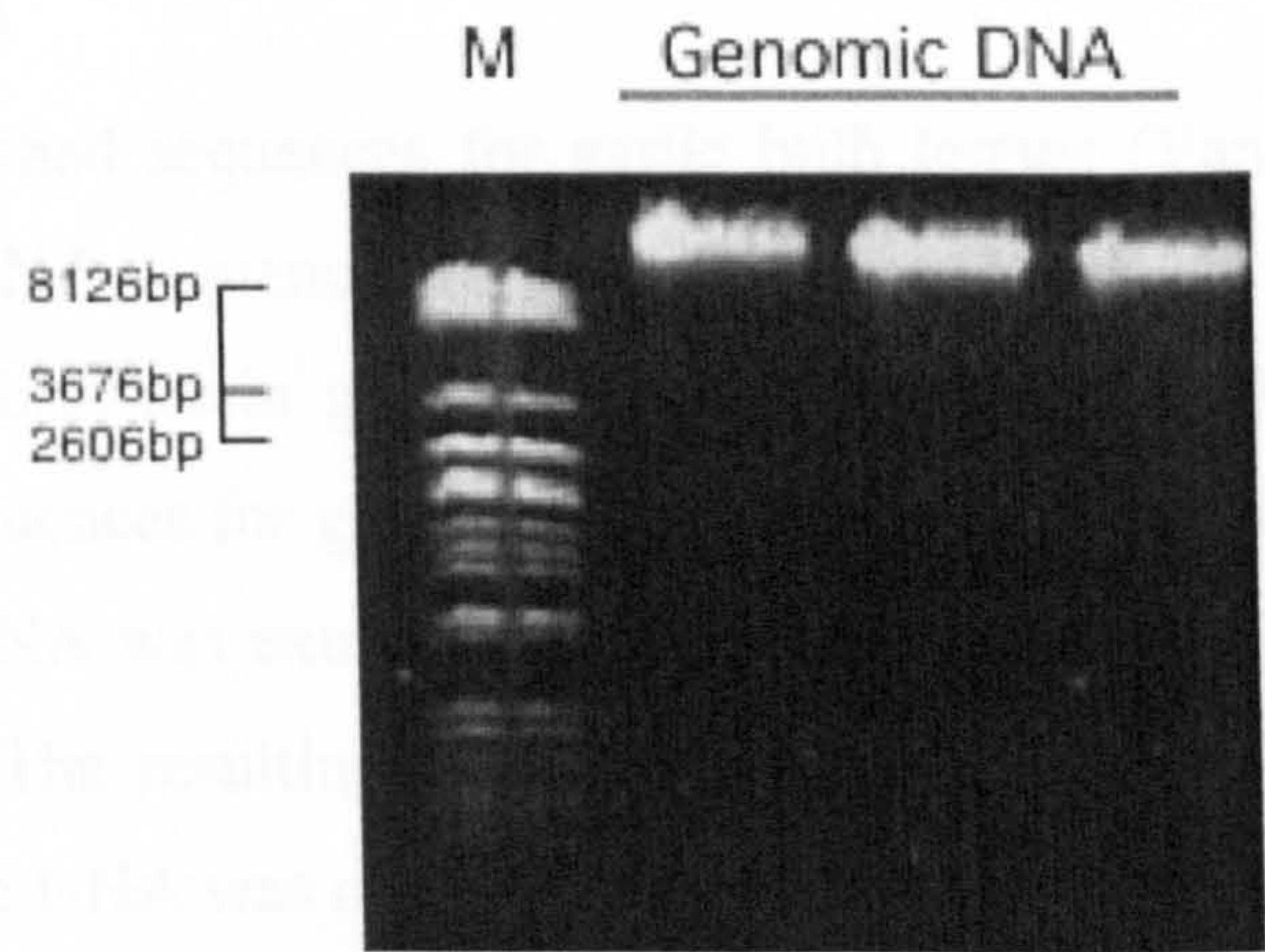


Figure 3.1 0.8% agarose gel showing genomic DNA isolated from a single *A. sativum* bulb. M = Eco471/ λ DNA molecular size markers.

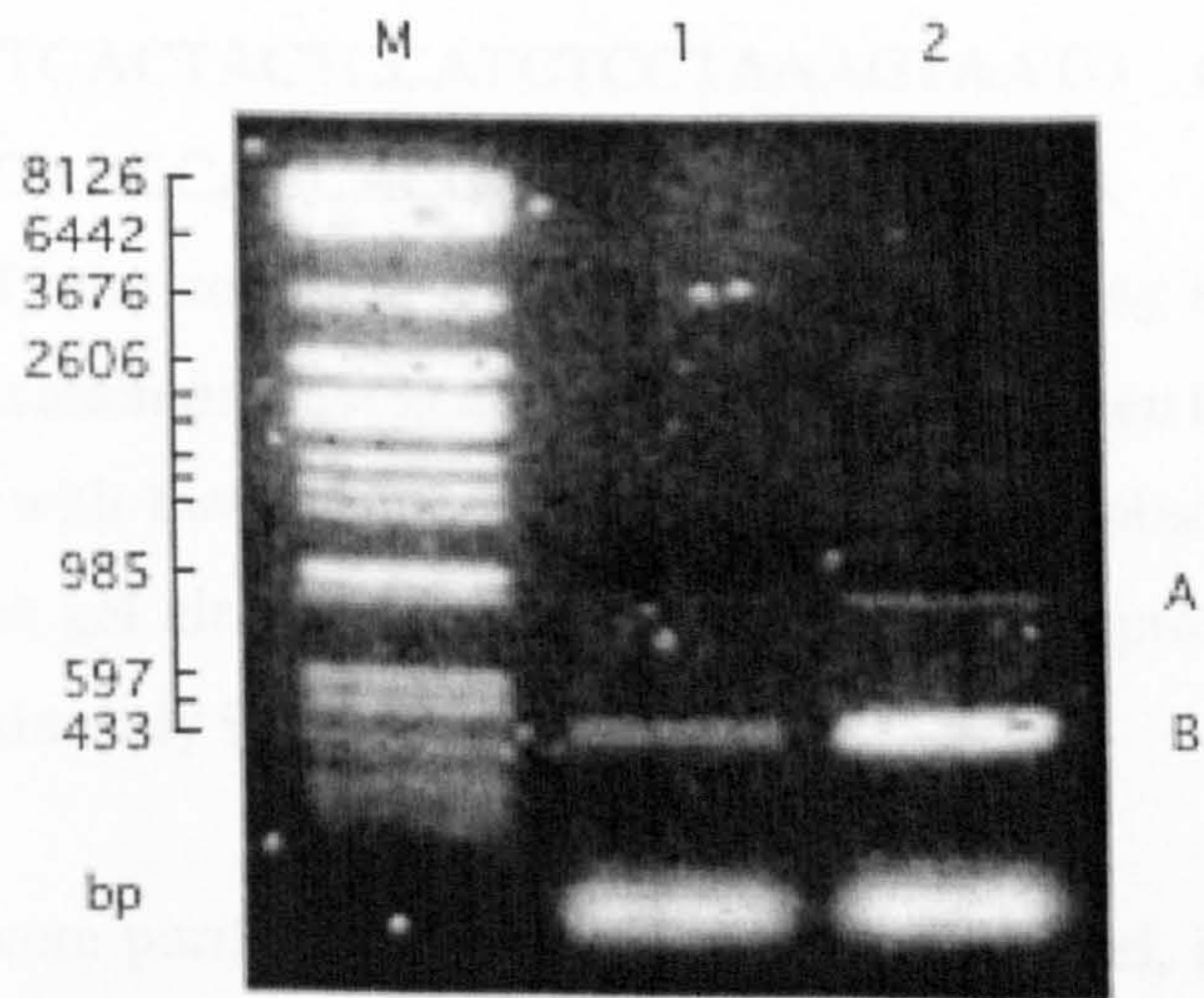


Figure 3.2 PCR of full-length ASAI (A) and ASAII (B) genes from *A. sativum* genomic DNA. M = Eco471/ λ DNA molecular size markers, 1 = 1 μ l genomic DNA template, 2 = 2 μ l genomic DNA template.

3.3.1 Isolation of Full Length ASAI and ASII Sequences from *Allium sativum* Genomic DNA

Analysis of the published sequences for garlic bulb lectins (Van Damme, E. J., *et al.*, 1992) and genomic DNA sequences for lectin genes from plants in general showed that none of the characterized lectin genes contained introns. This allowed amplification of full-length coding sequences for garlic lectins to be carried out using genomic DNA as a template. Genomic DNA was extracted from *A. sativum* bulbs using a commercial plant DNA extraction kit. The resulting DNA was assessed by agarose gel electrophoresis, which showed that the DNA was of high molecular weight, and did not contain fragments of < 10 kbp (fig. 3.1).

The PCR primers used to amplify lectin coding sequences were based on the published sequence for the homodimeric lectin isoform ASII 1 (accession number M85171):

Forward: AsLecF - ATGACTACTCCATCTCCTAAAGTAATG and

Reverse: AsLecR - TCAAGCAGCACCGGCCAGAGTGCCAA

These primers amplify the complete coding sequence, including start and stop codons. The predicted amplification product is a fragment of 467bp. When the purified *A. sativum* DNA was amplified with these primers, two fragments were observed as amplification products after agarose gel electrophoresis, a major band of approximately 470bp and a minor band of approximately 940bp (fig. 3.2).

The two fragments were purified by excision of bands from gel, isolation of DNA, and cloning in a specialised vector for PCR products (pCR2.1), using the TOPO TA cloning system. Clones resulting from this operation were characterised by DNA sequencing. The clone selected as representing the homodimeric garlic lectin, ASII, was identical to the published sequence for isoform ASII 3 (M85173.1) apart from four base differences which did not alter the encoded amino acids. The clone selected as representing the heterodimeric lectin, ASAI, was most similar to isoform 3 of ASAI (M85176.1), with approximately 93% similarity at the nucleotide level. Nucleotide and derived amino acid sequences of the ASAI and ASII clones used in this project are shown in figs. 3.3 and 3.4. Lectins in garlic are encoded by a multigene family (van Damme *et al.*, 1992), and multiple isoforms of the protein are predicted by the cDNA sequences present in the global database. In agreement with this observation, significant variability in the

sequences of clones encoding ASAI- and ASAII-like proteins, derived from amplification of genomic DNA, was observed, including several pseudogenes (both ASAI- and ASAII-like) containing stop codons midway through otherwise “correct” ORFs.

3.3.2 Preparation of Expression Constructs

Using the characterised clones encoding ASAI and ASAII as templates, coding sequences corresponding to mature lectin polypeptides were amplified for expression constructs, for use with the yeast expression vector pGAPZ α A. This shuttle vector contained the glyceraldehyde-3-phosphate (GAP) promoter, the *S. cerevisiae* α -factor secretory sequence, cloning site, myc-epitope, polyhistidine tag and alcohol oxidase 1 (AOX1) transcription terminator, a selectable marker gene (zeocin resistance) which functions in both *E. coli* and yeast, and a bacterial replication origin to allow propagation in *E. coli*. The primers were designed to amplify coding sequences extending from the first amino acid of the predicted mature polypeptide (using protein sequence data and the predictions of signal peptides given by the SignalP software), to the amino acid residue immediately before the stop codon. All oligonucleotides used for subcloning are shown in Table 3.1. The expression construct for ASAI was prepared using a PCR product amplified using the following primers:

Forward: ASAIpGAPz α A - CCGGAATTCAGGAACTTACTGACGAAC and

Reverse: ASAI/IIpGAPz α A – GCGGCGGCCGCAGCAGCACCGGCCAGAGT.

The primer sequences included a 5' EcoR I site and a 3' Not I restriction site to enable in-frame ligation into the pGAPz α A multiple cloning site.

The amplified PCR products were purified by agarose gel electrophoresis, and excised as bands from gel. DNA was purified from the gel slices and cloned into the subcloning vector pCR2.1 using the TOPO TA cloning system. The sequence of a selected clone was checked to confirm that no PCR errors were present, and the coding sequence insert was excised by digestion with EcoR I and Not I, and purified by agarose gel electrophoresis.

The coding sequence insert was ligated to pGAPZ α A vector DNA which had been digested with the same restriction enzymes, and recombinant plasmids were isolated by transformation of *E. coli* and selection on zeocin-containing media. The expression

<atg.....c.....g.....
1 actacttcatctcctaaactaatgagcatagccactgtagccgctatttctaaccatactggcatccacatgcatggc
M T T S S P K L M S I A T V A A I L T I L A S T C M A
P V
<----- Signal peptide ----->
.....a.....a..gg....a...aa....a.g.....g.
78 caggaacttactgacgaacggtgaaggactctacgctggccaatccctaaatgtggaaccctaccattttataatgcaag
R N L L T N G E G L Y A G Q S L N V E P Y H F I M Q
D N K
.....t.....t.....g.g...a.....g.gg...
158 aagactgcaacctcgtgctgtacgaccacagcactgccgtctgggccacaaactgacatccccggaaaaaagaatgc
E D C N L V L Y D H S T A V W A T N T D I P G K K E C
S G V H G
c.a..t...a.....g.....a....g.a....g.....a.....
238 agggcgggtgctgcagtctgatggcaactttgtggtctacgatgctgacggacgtcccctctgggccagtcacagtgtgaag
R A V L Q S D G N F V V Y D A D G R P L W A S H S V R
M A N H V N
.....c.....
318 agggaatggaaactatgtcctggtgcttcaagaggacgggaacggtgtcatttacggatctgatatttggtctactggtta
G N G N Y V L V L Q E D G N V V I Y G S D I W S T G
Q
.....a.....t.....g.....g.....
398 cctacgtaaaagctgcccgtggacctgtagttgtggctatgaatggtacgggttgatggaggctccctgattcgacccgta
T Y V K A A R G P V V V A M N G T V D G G S L I R P V
I G G
.c.....g.....a.....t..t.....
478 atcgtgaacaagaatcatgtccctgtcatccgaaagggttggcactctgtccaggaacatactgaggaacgacgaaggact
I V N K N H V P V I R K V G T L S R N I L R N D E G L
T D T
.....g.....c...t.....
558 gtacgcagggtcaatccctagatgtggaaccgtaccattttataatgcaagaagactgcaaccttggtgtgtacgaccaca
Y A G Q S L D V E P Y H F I M Q E D C N L V L Y D H
.....a.....g...a...t.....
638 gcaactgccgtctgggcctcaaactgacatccccggaaaaaaggatgcaaggcgggtgctgcagtctgatggcaacttt
S T A V W A S N T D I P G K K G C K A V L Q S D G N F
T R
.....
718 gtggtctacgatgctgaggggcggtccctctgggccagtcacagtgtgaagagggaatggaaactatgtcctggtgcttca
V V Y D A E G R S L W A S H S V R G N G N Y V L V L Q
.....a.....c..c.....t...a.....t.g.....
798 agaggacgggaacggtgtcatttacggatctgatatttggtcaactgatacctacagaaaaagtgctcgtggacctgtca
E D G N V V I Y G S D I W S T D T Y R K S A R G P V
I N V G
.....>
878 tccgaaagggttggcactctggccggtgctgcttgagttttaagatgtgtgttccatgttttaagtttgaataatgaataa

I R K V G T L A G A A *>

958 tgtgcgtgatgggttatgcactgtacgtatgcagctggtatgctgtatgtatattaagtcaccgtttgctggt

Figure 3.3 Sequence of 940bp PCR product clone (indicated by <.....> as limits) compared to cDNA encoding garlic lectin ASAI (isoform 3; M85176). Base differences are noted above the reference sequence. Changes in amino acid sequences resulting from base changes are shown below the reference amino acid sequence.

```
<.....
1 aaagaatgactactccatctcctaaagtaatgagcatagccactgtagccgctatttctaaccatatgggcatccacatgc
  <M T T P S P K V M S I A T V A A I L T I W A S T C
  <----- Signal peptide -----
.....t.....a.....C.....
81 atggccaggaacatactgatgaacgacgaaggactgtacgcaggccaatccttagatgtggaaccgtaccatcttataat
  M A R N I L M N E D G L Y A G Q S L D V E P Y H L I M
----->
.....C.....
161 gcaggaagactgcaacttcgtgctgtacgaccacagcactgccgtctggacctcaaacactgacatccccggaaaaaag
  Q E D C N L V L Y D H S T A V W T S N T D I P G K K
-
.....
241 gatgcaaggcggtgctgcagtcctgatggcaactttgtggtctacgatgccgagggacgttccctctgggccagtcacagt
  G C K A V L Q S D G N F V V Y D A E G R S L W A S H S
.....
321 gtaagaggggaatggaaactatgtcctgggtgcttcaagaggacggtaacgttgtcatttacggatctgatatttggtctac
  V R G N G N Y V L V L Q E D G N V V I Y G S D I W S T
.....>
401 taatacctacaaaggaagtgcctcgtggacctgtcatccgaaagggttggcactctggccggtgctgcttgagttttaggat
  N T Y K G S A R G P V I R K V G T L A G A A *>
481 gtgtgttccatgtttttaagtttgaataatgaataatgaataatgtgcgtgatggtgatgcactgtacttatgtcaactgt
561 tatgctgcattatatatattaagcacagtttgctggtta
```

Figure 3.4 Sequence of 470bp PCR product clone (indicated by <.....> as limits) compared to cDNA encoding garlic lectin ASAI (isoform 3; M85173). Base differences are noted above the reference sequence.

Name	Length	Sequence	Translation	Description
AsLecF	27mer	ATGACTACTCCATCTCCTAAAGTAATG	MTTPSPKVM	5' ASAIL signal peptide start
AsLecR	26mer	TCAAGCAGCACCGGCCAGAGTGCCAA	GTLAGAA.	3' ASAIL 3' end
5' ASAIpGAP α A	27mer	CCGGAAATTCAGGAACTTACTGACGAAC	RNLLTN	5' ASAI mature peptide start + EcoRI site
3' ASAILpGAP α A	29mer	GCGGCGGCCGAGCAGCACCGGCCAGAGT	TLAGAA	3' ASAI and ASAIL mature peptide end + NotI site
5' ASAILpGAP α A	27mer	CCGGAAATTCAGGAAACATACTGATGAAT	RNILMN	5' ASAIL mature peptide start + EcoRI site

Table 3.1 Oligonucleotide sequences for ASAI and ASAIL expression in pGAP α A. Italic indicates a restriction site and underline indicates where the oligonucleotide is complementary to the template sequence.

construct in a selected clone was re-checked by DNA sequencing to confirm correct assembly. The predicted protein product from the ASAI expression construct is shown in fig. 3.5.

A similar strategy was used to assemble the expression construct for ASAII. Again, the oligonucleotide primers for amplification of the ASAII mature peptide sequence were designed containing a 5' EcoR1 and a 3' Not1 restriction site. The oligonucleotide sequences were as follows:

Forward: ASAIIpGAPzαA – 5' – CCGGAATTCAGGAACATACTGATGAAT – 3'

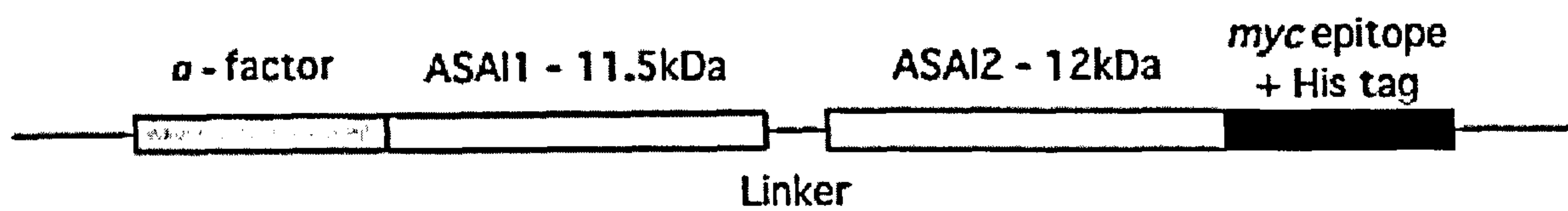
and

Reverse: ASAI/IIpGAPzαA – 5' – ACTCTGGCCGGTGCTGCTGCGGCCGCCGC – 3'

The predicted protein sequence from the ASAII expression construct is shown in fig. 3.6. Amino acid sequence predictions of recombinant ASAI and ASAII are shown in fig 3.7.

3.3.3 Transformation of Expression Constructs into *Pichia pastoris*

The constructs described above were transformed into *Pichia pastoris* cells (SMD1168 and X-33 strains) by chemical transformation. Transformants were selected by plating on zeocin-containing medium (100 µg/ml Zeocin, YPGZ). Typically, between 10-20 zeocin-resistant colonies were present after 3 days of incubation at 30 °C from a transformation using 10 µg linearised pGAPZα DNA. To screen for expression of recombinant protein, individual colonies were grown in 10 ml of liquid media for a further 3 days. Culture supernatant was assayed by precipitation of total protein with ammonium sulphate, followed by analysis of the pellet by SDS-PAGE and western blotting, using anti-ASA or anti-myc antibodies. Typical blots from these screens for ASAI and ASAII are shown in figs 3.8 and 3.9. The colony with the highest expression level in this screen was chosen for use in fermentation.



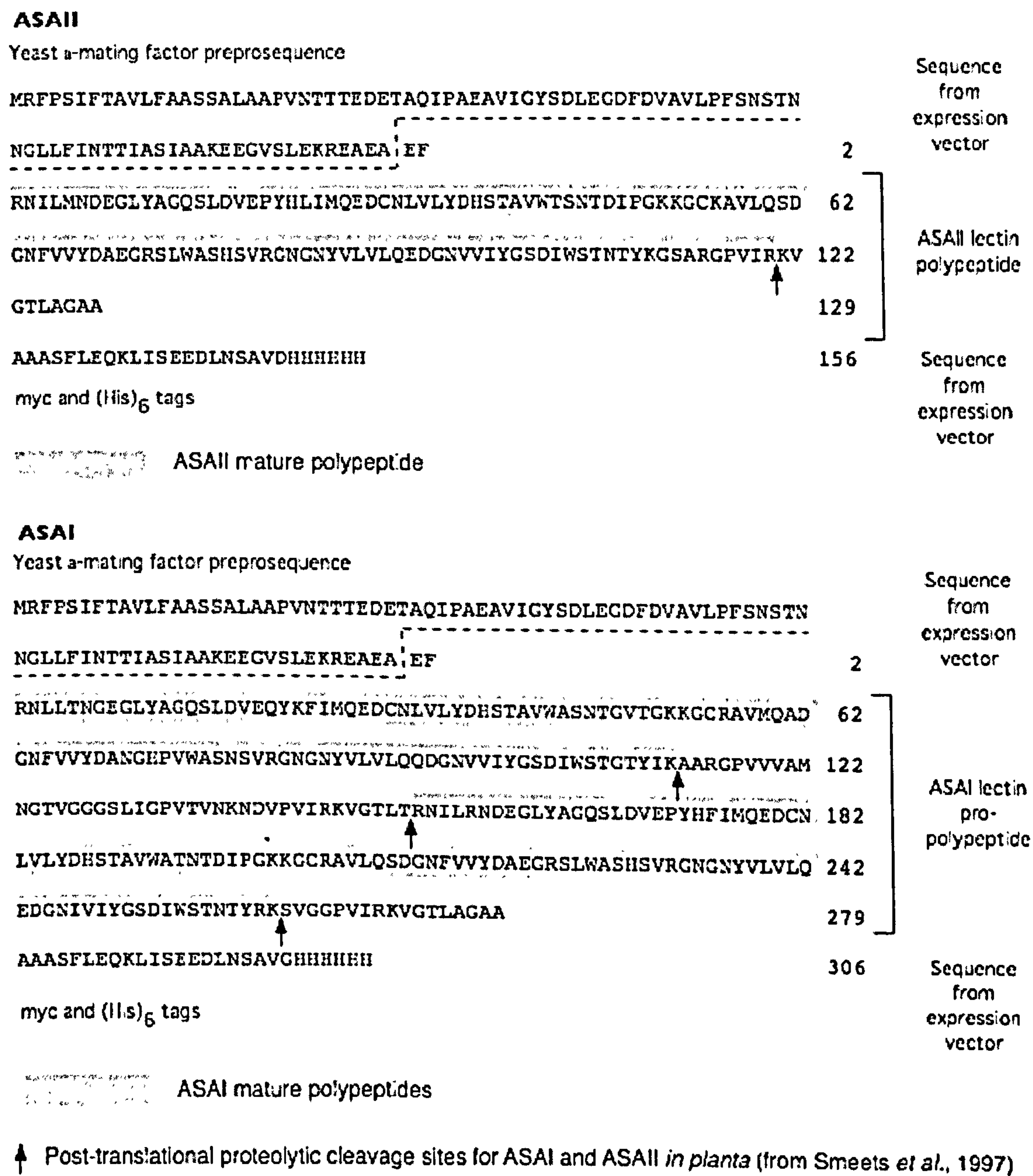
Heterodimer - 940bp, 306 amino acids

Figure 3.5 ASAIpGAPzαA expression construct



Homodimer - 470bp, 154 amino acids

Figure 3.6 ASAIIpGAPzαA expression construct



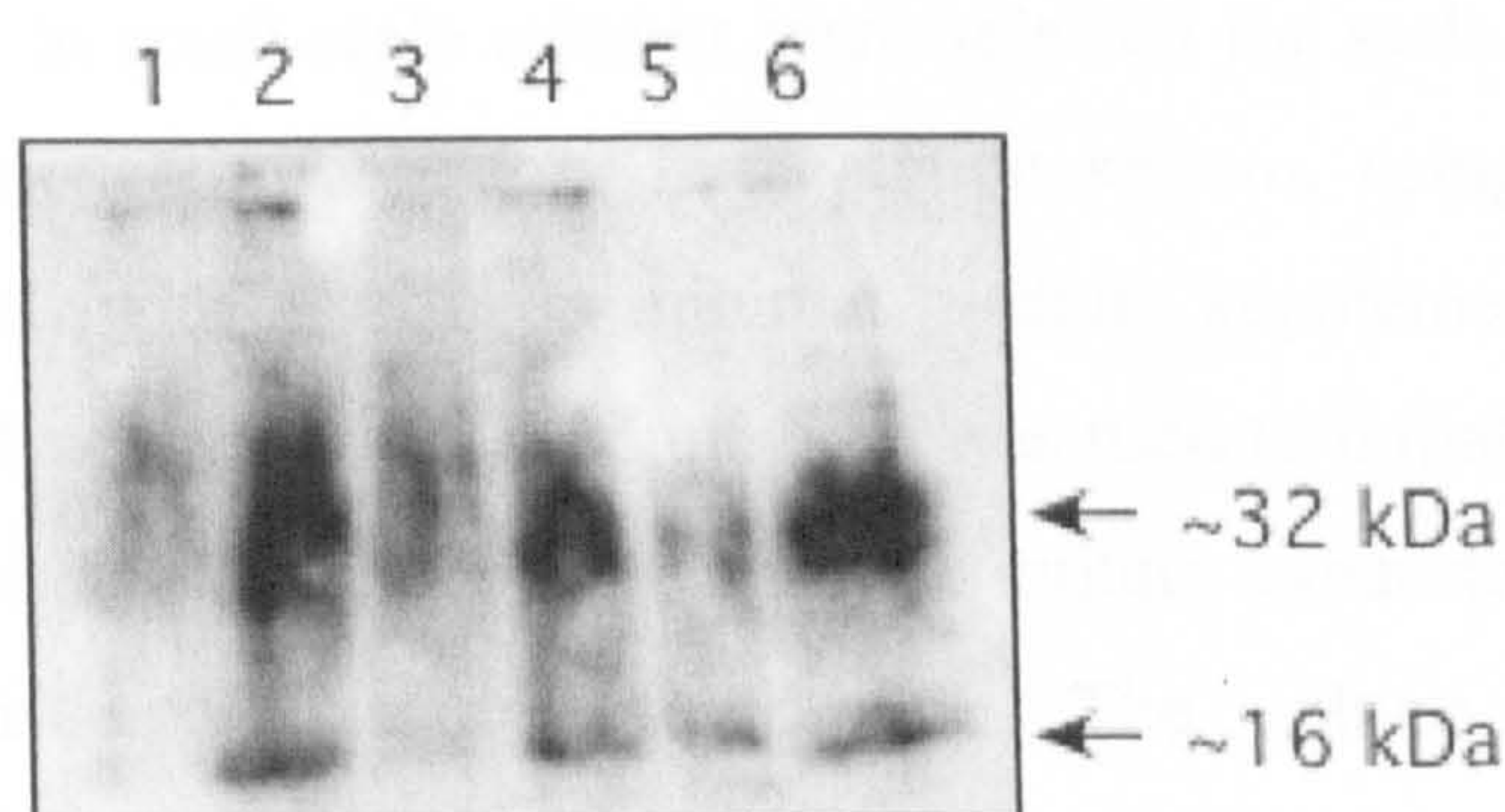


Figure 3.8 Protein expression screen of ASAI in *P. pastoris* strain SMD1168. Samples of culture supernatant from small-scale shake-flask cultures were taken after 3 days growth, precipitated with ammonium sulphate, and samples of the resulting pellet were analysed by SDS-PAGE and Western blotting. An anti-myc epitope antibody was used for detection.

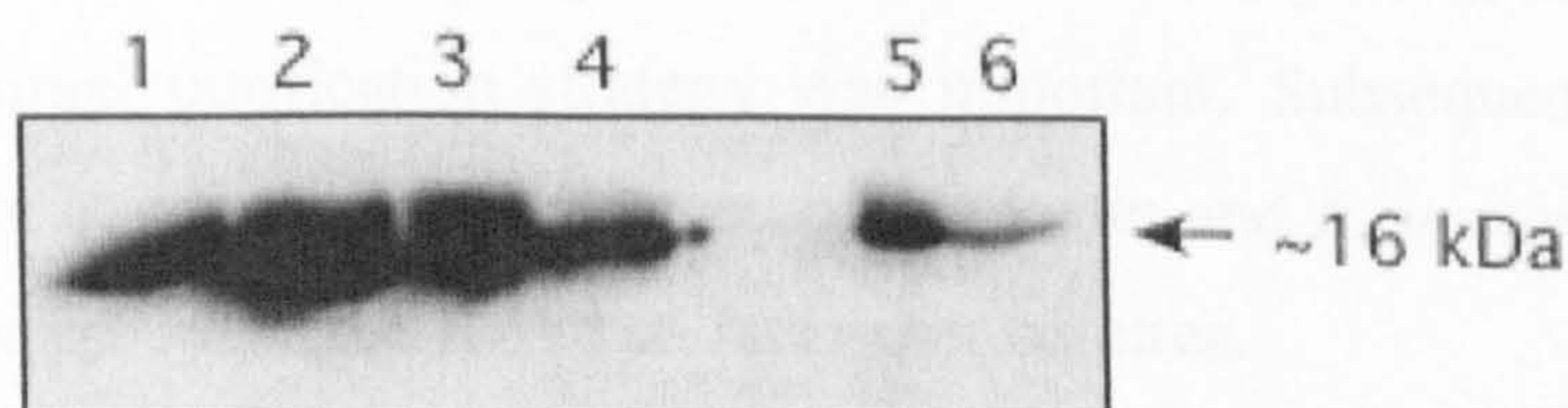


Figure 3.9 Protein expression screen of ASAII in *P. pastoris* strain SMD1168, carried out as described above. Detected using anti-myc epitope antibody.

3.3.4 Scale-up of Production of Recombinant Garlic Lectins in *Pichia pastoris*

Clones of the expression constructs for ASAI and ASAII that had been selected as the highest expressers in small-scale cultures were selected for scale-up of production. The scale-up employed a laboratory fermentation system, using a 2.5 l or 7.5 l fermentation vessel, with growth on minimal medium supplemented by a glycerol feed. A starter culture grown in YPGZ medium was used to initiate the fermentation. Fermentation typically lasted 4-5 days, until the culture exhausted its glycerol feed and oxygen consumption started to decline. The culture supernatant from fermentation was harvested and centrifuged to remove cell debris. Overall expression levels of ASAI and ASAII were similar, with 10-20 mg of purified recombinant protein recovered per litre of fermenter culture supernatant. Fermentation produced a significantly higher yield than was achievable in shake-flask cultures due to the far greater cell densities achieved and constant optimization of conditions. In YPG shake flask cultures containing 2 l of medium in a 5 l baffled conical flask, typical yields were approximately 1 mg/l after 5 days growth.

3.3 Purification and Characterisation of Recombinant Garlic

Lectins expressed in *P. pastoris* strain X-33

P. pastoris culture supernatant is amenable to a range of protein purification techniques. In order to produce large yields of recombinant protein for insect feeding studies, an optimal purification strategy was important. Subsequently, a range of procedures were attempted for each recombinant lectin and *P. pastoris* strain, which were applicable to both shake-flask and fermenter cultures.

Metal affinity chromatography, exploiting the (his)₆ tag added to the C-terminal end of the expression construct was attempted. However, binding of the recombinant lectins expressed in the wild-type yeast strain X-33 to a Ni²⁺ column was very poor, and only a very small proportion of the lectin could be eluted from the column with high concentrations of imidazole buffer, most lectin being present in the flow-through fraction (result not shown). Analysis by Western blotting using anti-myc and anti-(his)₆ antibodies showed that the lectin in the flow-through fraction was not

immunoreactive, suggesting that the C-terminal tag sequences had been lost by proteolysis, and explaining the failure to bind. This method was not pursued further. Purification by mannan-affinity chromatography did not have adequate binding capacity to purify all recombinant lectin from culture supernatant and approximately 50% of recombinant protein was recovered from column flow-through. The matrix could however be used to show *in vitro* mannose-binding activity as demonstrated later in this chapter. Hydrophobic interaction chromatography on columns of phenyl-Sepharose was a viable purification strategy for ASAI, with the lectin binding to the column in buffer containing 2 M salt, and being eluted by water. This technique was used to purify polypeptide 2 (C-terminal) from ASAI, with a separate “clean-up” step by ion-exchange chromatography on a Q-ion exchanger being incorporated after the hydrophobic interaction chromatography step (fig. 3.10). However, the columns routinely failed to bind ASAII, despite various modifications to the procedure.

Ion-exchange chromatography was selected for large-scale purification of lectins from culture supernatants, as a robust, easily scalable method which could be used with large volumes. Both lectins could be purified using cation-exchange chromatography on S-Sepharose at pH 4.0 in 50 mM sodium acetate buffer, with the binding and washing steps being carried out at 4 °C. For binding to occur it was necessary to dilute the initial fermenter culture supernatant 10-fold into 50 mM sodium acetate, in order to lower the concentration of sodium chloride in the culture supernatant (approximately 1.0 M) sufficiently to enable binding to the matrix. After binding and washing the column, the lectins were eluted from the column at room temperature with a salt gradient, eluting in the range 0.15 - 0.3 M NaCl. A typical elution profile is shown (fig 3.11). After dialysis and lyophilization of the peak fractions from initial purification, a further clean-up purification step was carried out by gel filtration on a column of Sephacryl S-200, which removed the high-molecular weight contaminating yeast proteins present after the primary purification. A typical column profile is shown (fig. 3.12). After gel-filtration chromatography, recombinant lectins were free of contaminating proteins (>95% homogenous by SDS-PAGE) and were again dialysed and lyophilized. The resulting lyophilate typically contained 40-60% dry weight recombinant protein, the remaining dry weight being accounted for by yeast carbohydrate. Despite carrying out column loading and washing at 4 °C, the extended

times required to load the large volumes involved (approximately 3 days) meant that proteolysis of recombinant lectins could occur. However, despite the observation that specific proteolysis of lectins expressed in *P. pastoris* strain X-33 did occur during expression there was no evidence for non-specific general proteolysis (shown by the presence of “smears” of degraded protein on gels) during purification.

3.3.1 SDS-PAGE of recombinant ASII produced in *P. pastoris* X-33

Analysis of purified recombinant proteins produced in the wild-type yeast strain X-33 by SDS-PAGE showed a band of the expected molecular weight for ASII. Results are shown in fig. 3.12. ASII (homodimeric lectin) gave a single band at approximately 14.5 kDa. The recombinant ASII has a higher molecular weight than that reported for ASII purified from plant tissue. However, estimation of molecular weight by SDS-PAGE is not accurate in this low molecular weight range unless gels containing urea are used. The purified lectin did not react with anti-myc or anti-(his)₆ antibodies (result not shown). These data establish that proteolysis of the recombinant protein product is taking place during expression. As well as the expected cleavage of the N-terminal α -factor pro-sequence, cleavage of the C-terminal extensions containing the epitope tags must be taking place for ASII. Proteolysis of the initially expressed polypeptide, after removal of the α -factor pro-sequence, is most likely to be a result of action of the yeast extracellular aspartic proteinase encoded by the *Pep4* gene. This result will be considered in more detail later in this chapter.

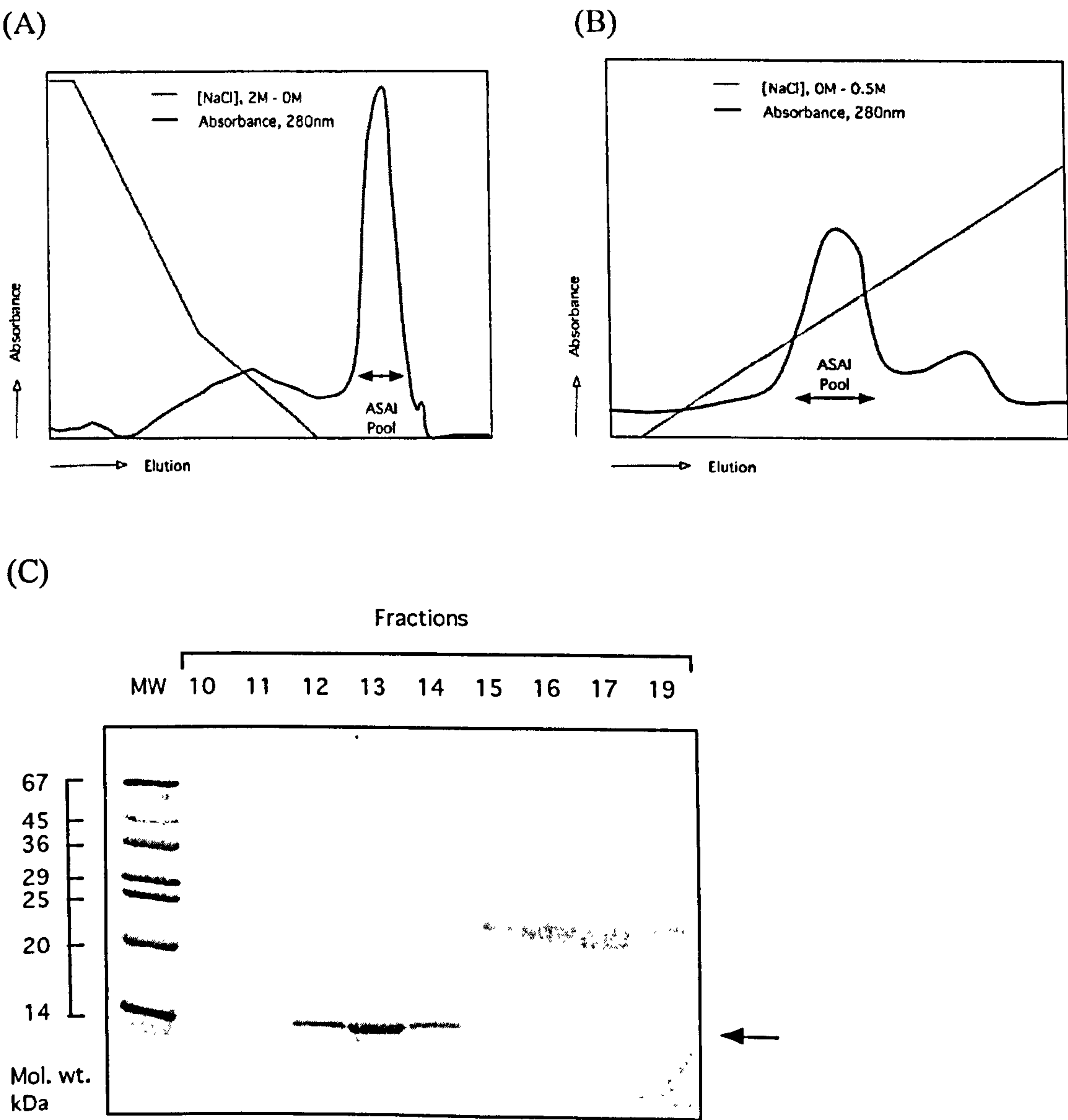
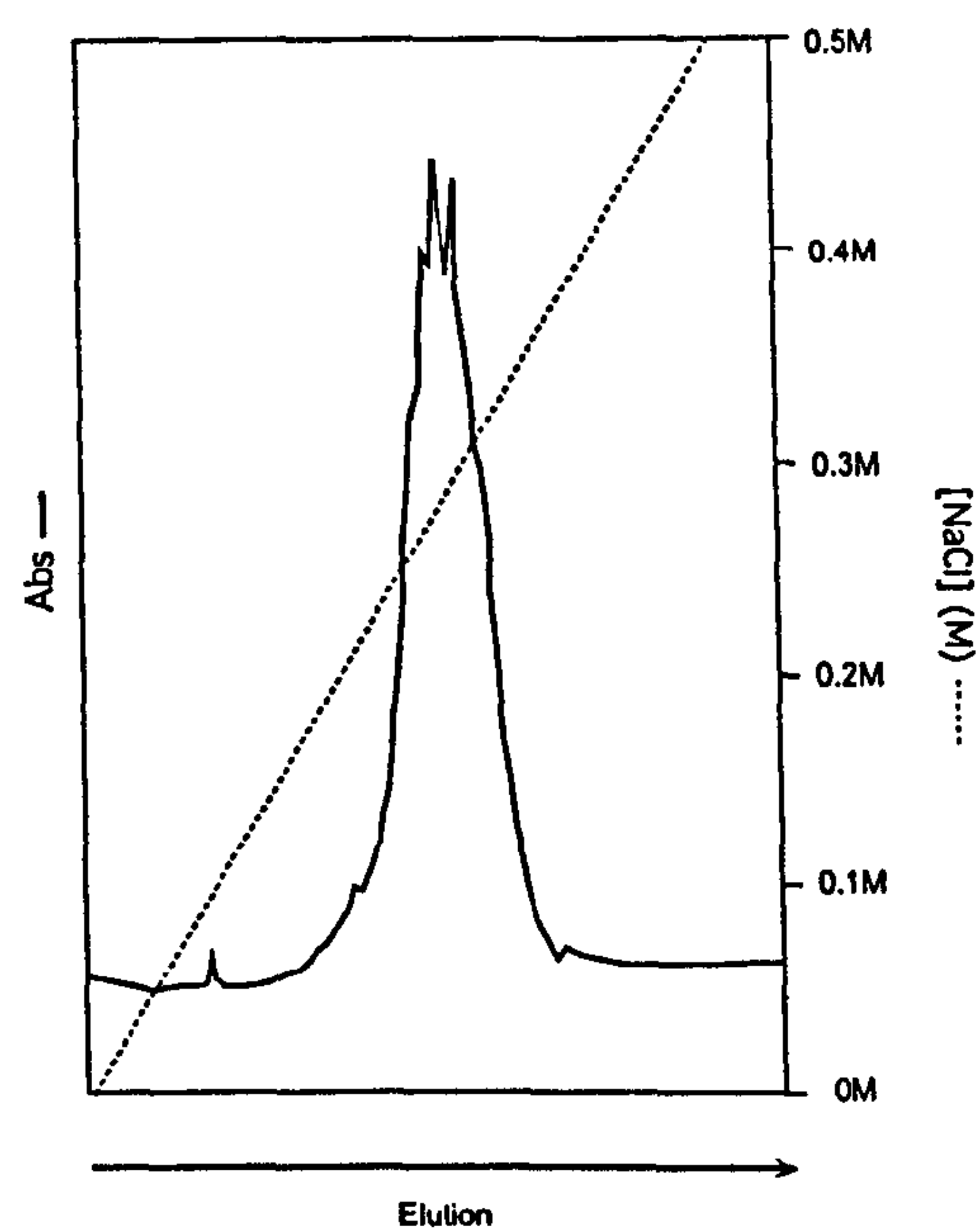


Figure 3.10 Purification of recombinant ASAI (polypeptide 2), produced in wild-type *Pichia pastoris* (strain X-33). Protein was purified from *P. pastoris* culture supernatant by hydrophobic interaction chromatography on a phenyl-Sepharose column (A). Peak fractions were pooled and further purified by ion exchange chromatography on a HiTrapQ column (B) in 20 mM diaminopropane buffer, pH 9.0. The column was eluted with a linear gradient of NaCl from 0 M (fraction 3) to 0.5 M (fraction 28). Fractions were analysed by SDS-PAGE (C). Arrow indicates ASAI (polypeptide 2). MW = molecular weight markers.

(A) ASAI



(B) ASAII

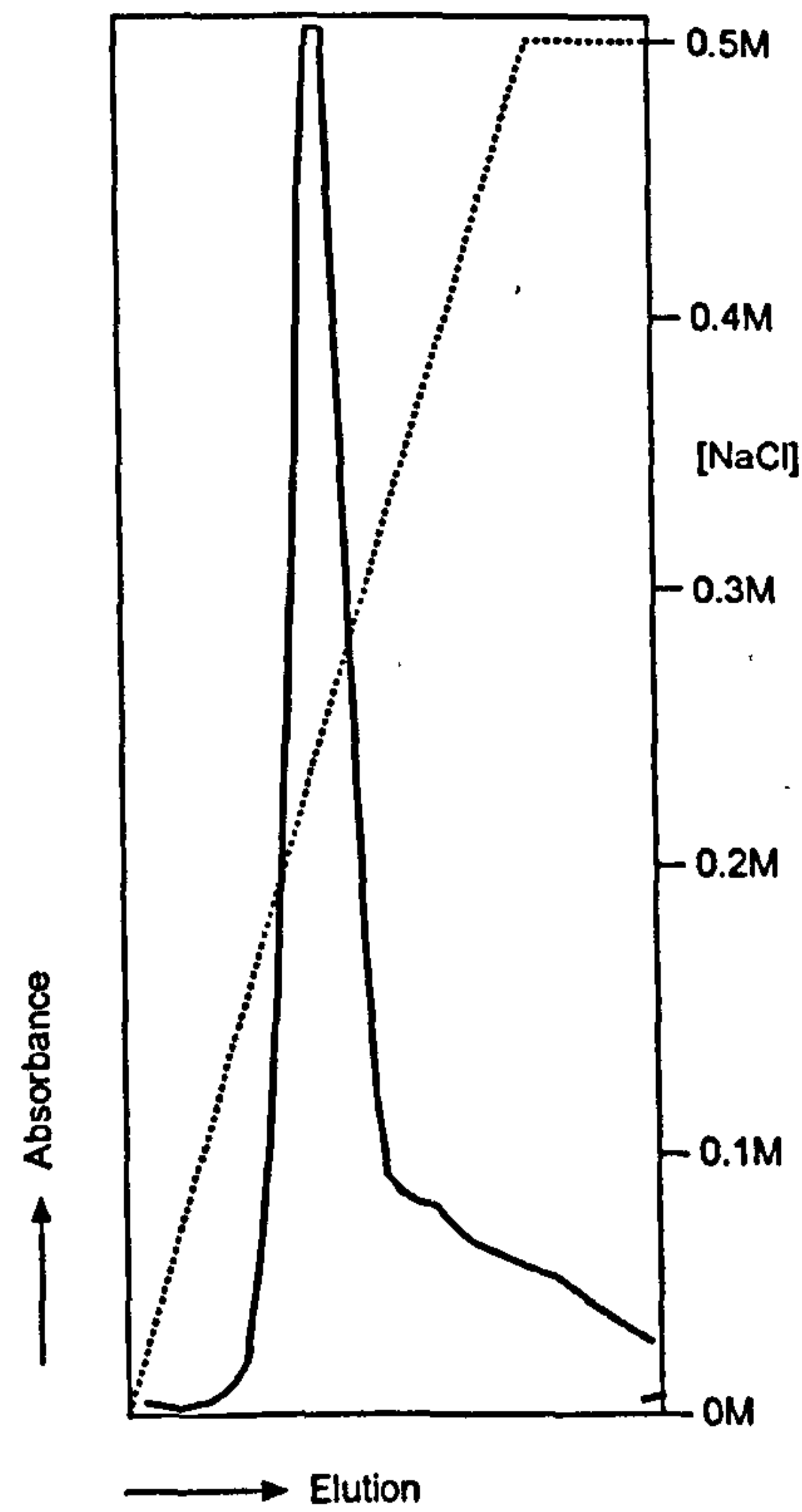


Figure 3.11 Purification of recombinant garlic lectins from yeast culture supernatant by ion-exchange chromatography on S-Sepharose at pH 4.0.

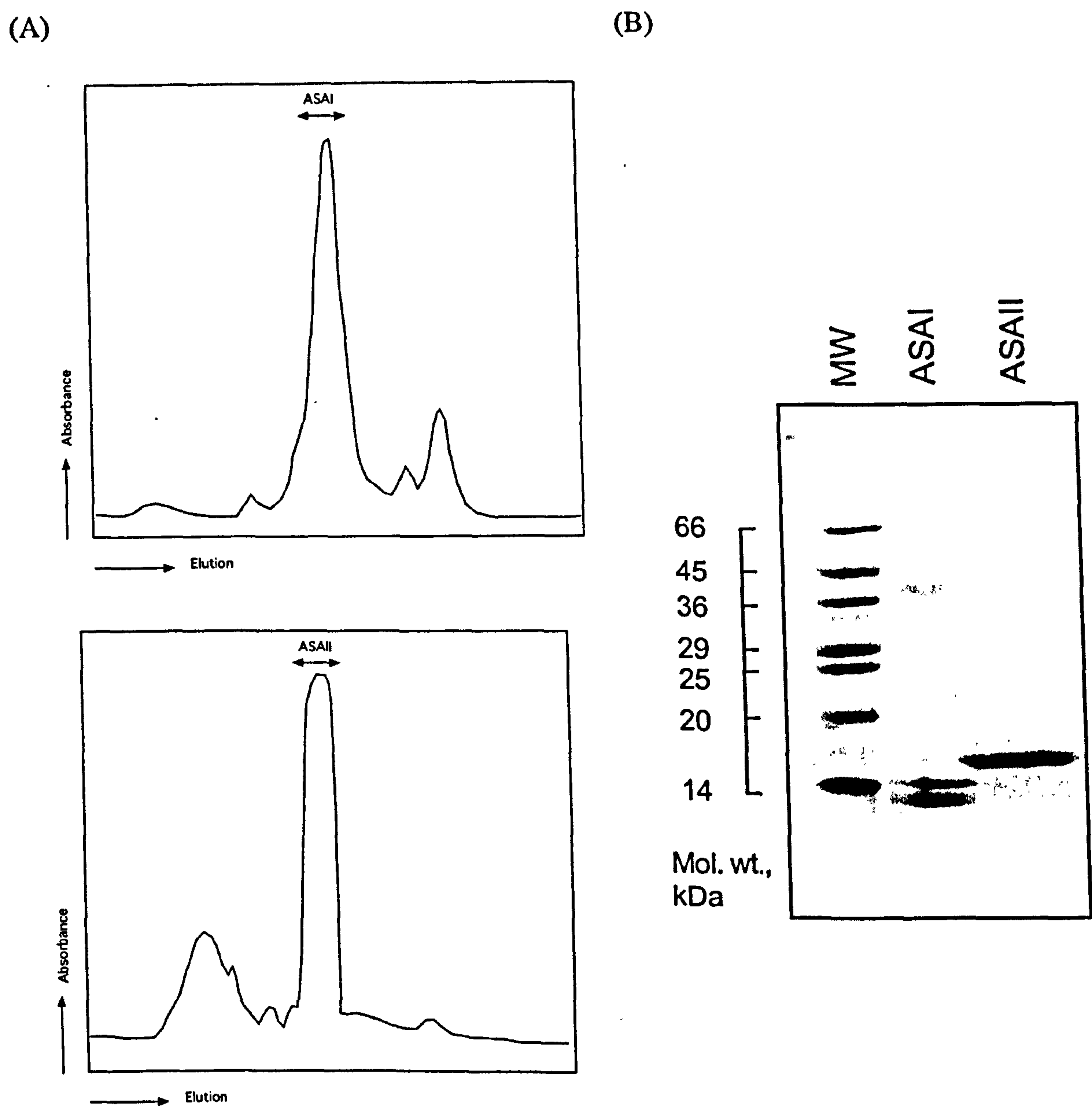


Figure 3.12 (A) “Clean-up” purification profiles for ASAI and ASAIL. Gel filtration was carried out on a column of Sephacryl S-200. (B) SDS-PAGE of samples of ASAI and ASAIL expressed as recombinant proteins in *P. pastoris* strain X-33 (wild type) after purification.

3.3.2 ASAI produced in *P. pastoris* strain X-33 shows variable polypeptide band patterns on SDS-PAGE

The degree of proteolysis observed in different preparations of ASAI produced in *P. pastoris* strain X-33 was variable, depending on fermentation conditions and non-reproducible factors. While some samples showed complete proteolysis to the polypeptides of 13.5 kDa and 14 kDa shown in fig. 3.12, others showed the presence of higher molecular weight bands indicative of incomplete proteolytic cleavage to polypeptides similar in size to those found in the protein found *in planta*. For example, in a sample of recombinant ASAI showing incomplete proteolysis, protein bands at approximately 13.5 kDa, 14 kDa, 16 kDa, 32 kDa and a diffuse band in the range 40-65 kDa were present (figs. 3.17 and 3.18). Of these bands, the 32 kDa band is most likely to represent the initial predicted product after removal of the α -factor pro-sequence (predicted molecular weight 33114 Da). The identity of the other bands is considered below.

3.3.3 Analytical Gel Filtration to determine ASAII oligomerization

In order to determine if recombinant ASAII produced in *P. pastoris* formed dimeric molecules, as is the case for the lectin purified from plant tissues, analytical gel-filtration was carried out on a Superdex 200 column. The column was calibrated with standard proteins BSA, ovalbumin, and soybean Kunitz trypsin inhibitor. The retention volumes of recombinant ASAII produced in *P. pastoris* strains X-33 and SMD1168 (see below) were measured (fig. 3.13). Both proteins eluted as a single major peak, with an estimated molecular weight of $26,500 \pm 1000$ Da; the elution volume of the protein produced in X-33 was slightly greater than that produced in SMD1168, but the difference was within experimental error for the system. This result shows that the recombinant protein is able to form dimeric molecules containing two polypeptides.

3.3.4 *in vitro* Haemagglutination Activity of Recombinant Garlic Lectins

Both recombinant ASAI and ASAII expressed in X-33 were able to agglutinate fresh rabbit erythrocytes. It has been reported that native ASAI has approximately 10-fold higher agglutination activity than native ASAII (Van Damme, E. J. M., *et al.*, 1998). This was not observed in the case of recombinant garlic lectins, both ASAI and ASAII X-33 agglutinated rabbit erythrocytes down to a dilution of approximately 0.5 mg/ml, compared to approximately 0.05 mg/ml in the case of a native GNA standard (fig. 3.14). Low agglutination activity of garlic lectin on rabbit erythrocytes has been observed for protein purified from plant tissues. The demonstration of biological activity for the recombinant proteins is in agreement with earlier observations that binding to mannan-agarose columns could be observed with both ASAI and ASAII from yeast culture supernatants (see above). Experiments carried out with purified proteins confirmed their ability to bind to mannan-agarose, although complete binding was not observed under the conditions employed (data not presented).

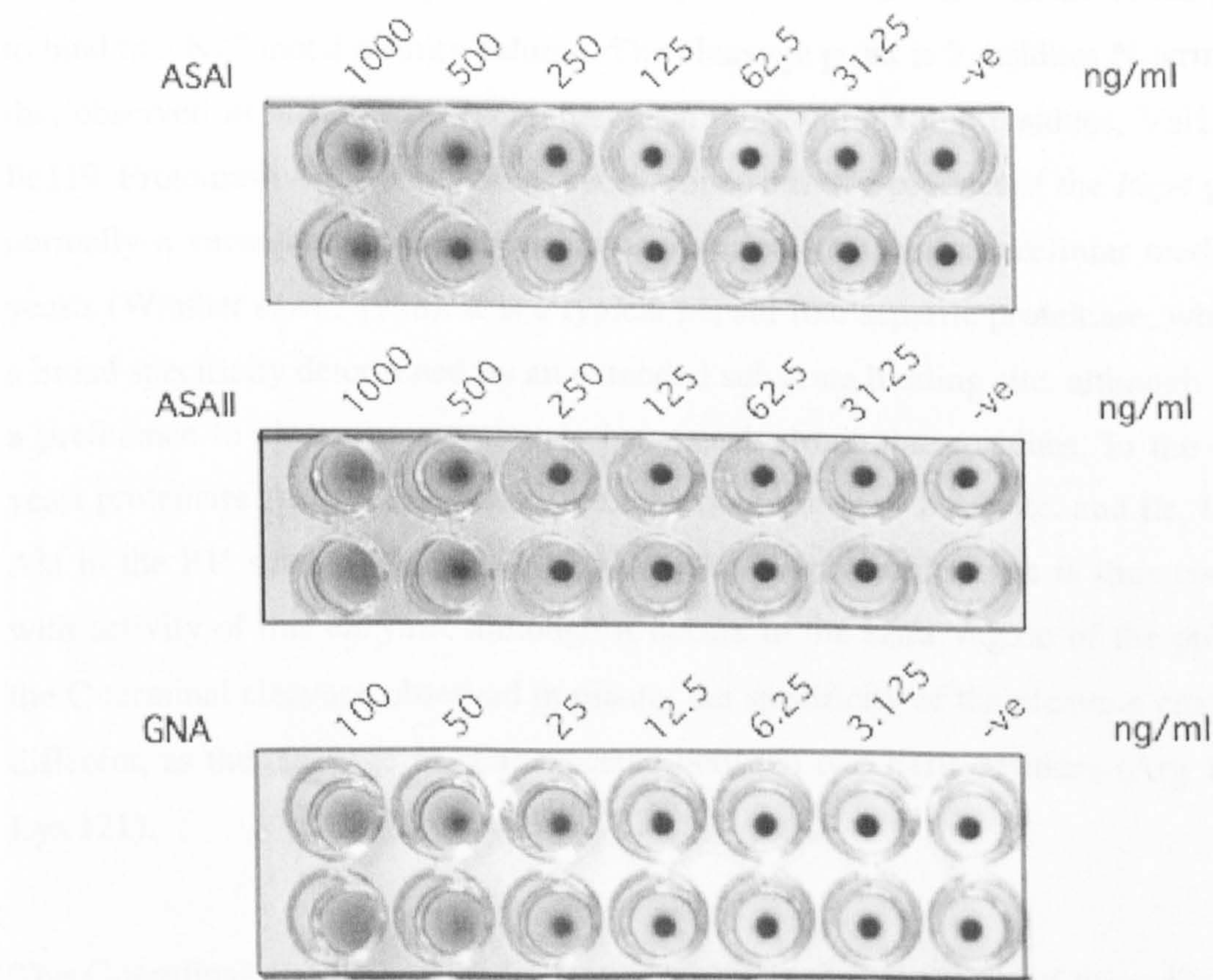
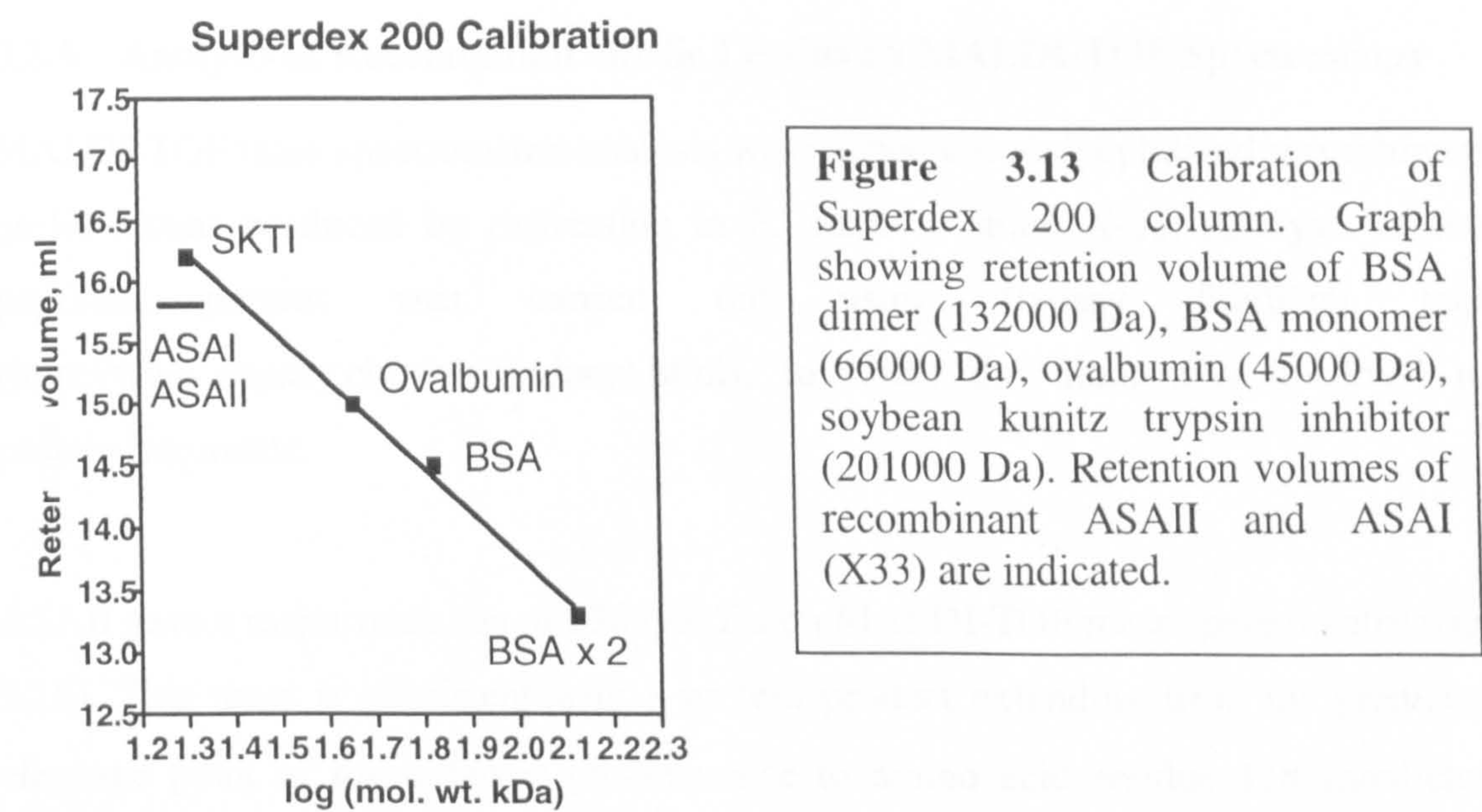


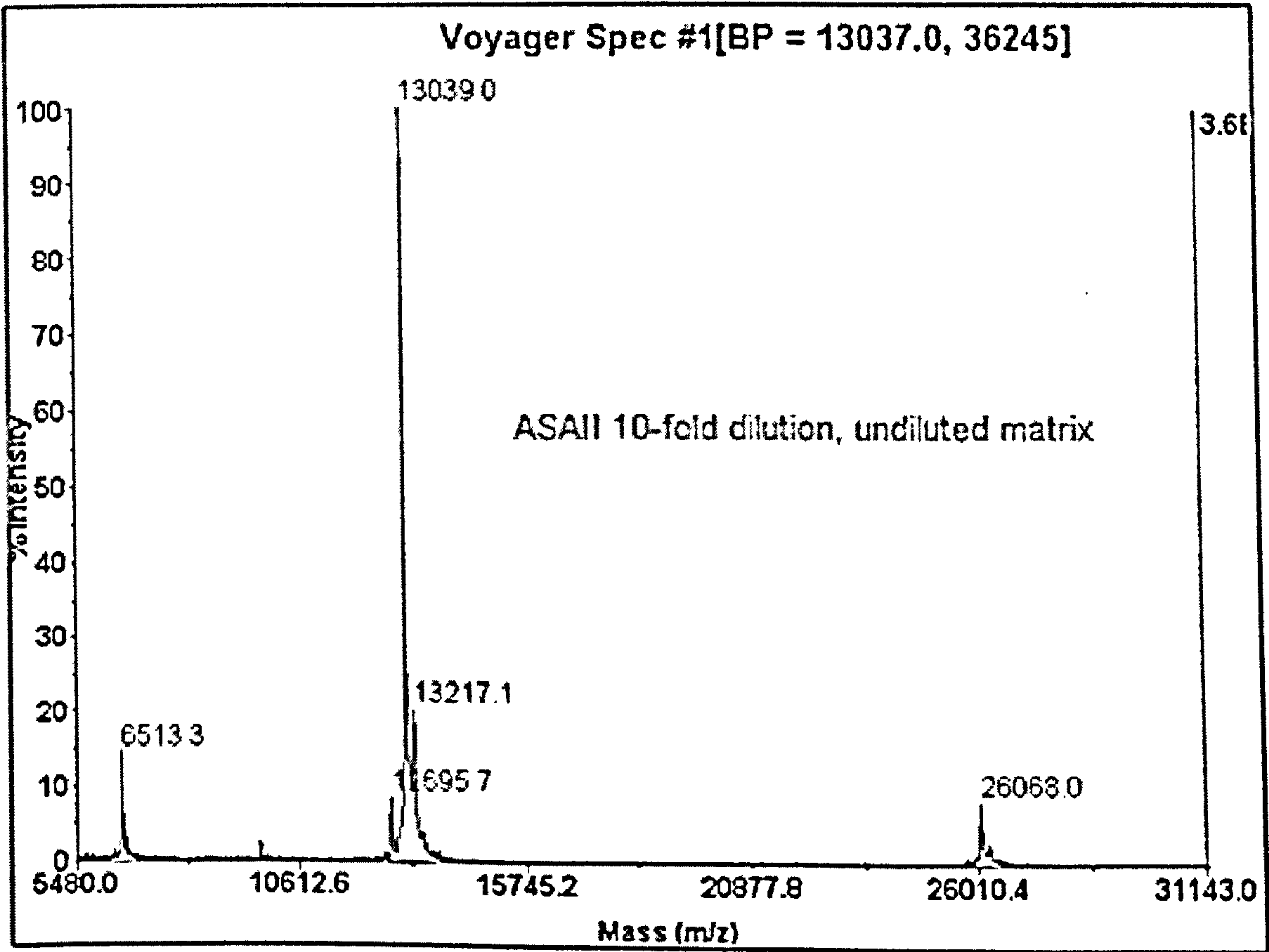
Figure 3.14 Haemagglutination assays demonstrating *in vitro* biological activity of recombinant ASAI, ASAII and commercially available GNA standard. Values indicate ng/ml of lectin.

3.3.5 Analysis of Recombinant Garlic Lectins by MALDI-TOF Spectroscopy

MALDI-TOF mass spectrometric analysis was carried out on lyophilized recombinant garlic lectins produced by expression in *P. pastoris* strain X-33. Analysis of the peptides present was carried out using Expasy FindPept tool (<http://www.expasy.ch/tools/findpept.html>), to relate the mass ions observed to protein fragments.

ASAII gave a major mass ion of 13039.0 Da on MALDI-TOF mass spectrometry (fig. 3.15). This mass is consistent with a protein product extending from the predicted cleavage point of the α -factor pro-sequence to amino acid residue 118 (predicted mass 13039.3). C-terminal cleavage to remove the myc and (his)₆ “tag” sequences from the recombinant protein has thus occurred, as suggested by the lack of antigenicity towards anti-myc and anti-(his)₆ antibodies, and the failure of the protein to bind to a Ni²⁺ metal affinity column. The cleavage point is 2 residues N-terminal to that observed *in planta*, and occurs between two hydrophobic residues, Val118 and Ile119. Proteinase A (also known as saccharopepsin), the product of the *Pep4* gene, is normally a vacuolar proteinase, but is also secreted to the extracellular medium by yeasts (Winther *et al.*, 1998). It is a typical pepsin-like aspartic proteinase, which has a broad specificity determined by an extended substrate binding site, although there is a preference to cleave peptide bonds between hydrophobic residues. In the case of yeast proteinase A, Phe, Leu and Glu are favoured in the P1 subsite, and Ile, Leu and Ala in the P1' subsite (Dreyer, T., 1989). The observed cleavage is thus consistent with activity of this enzyme; although it occurs in the same region of the protein as the C-terminal cleavage observed in *planta*, the specificity of the cleaving enzymes is different, as the cleavage *in planta* occurs between two basic residues (Arg 120 and Lys 121).

The C-terminal cleavage of garlic lectin ASAII is thus determined primarily by the structure of the protein, rather than the specificity of the cleaving enzyme; cleavage occurs in the C-terminal region because this is exposed on the surface of the protein, and does not occur in the core region of the protein which has a more compact structure composed of β -sheets.



ASAII

Yeast α -mating factor preprosequence

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPFSNSTN

NGLLFINTTIASIAAKEEGVSLEKREAEA↑EF

RNILMNDEGLYAGQSLDVEFYHLIMQEDCNLVLYDHSTAVWTSNTDIPGKKGCKAVLQSD

GNFVVYDAEGRSLWASHSVRCNGNYVLVLQEDGNVVIYGSDIWSNTNTYKGSARGPVIRKV

GTLAGAA

AAASFLEQKLISEEDLNSAVDHHHHHHH

2

62

122

129

156

Sequence from expression vector

ASAII lectin polypeptide

Sequence from expression vector

ASAII mature polypeptide

Figure 3.15 MALDI-TOF mass spectrometric analysis of recombinant ASAII (produced in *P. pastoris* strain X-33). Sequence diagram shows proteolytic cleavage points on the predicted product giving a mass ion consistent with the major peak (open-headed arrows). Predicted mass of the protein = 13039.3, experimentally determined mass = 13039.0 Da.

The minor peaks at 12695.7 Da and 13217.1 Da observed in the mass spectrum of ASAII could not be unambiguously assigned to specific protein fragments. The peak at 6513.3 Da is putatively a doubly charged ion from the major peak at 13039, and the peak at 26068 Da is putatively a dimer of the 13039 species.

The mass-ion spectrum for recombinant ASAI (expressed in *P. pastoris* strain X-33) obtained by MALDI-TOF mass spectrometry was complex. Two major peaks, potentially corresponding to the two polypeptides observed on SDS-PAGE in this protein preparation were observed at 12863.0 Da and 13758.7 Da (fig. 3.16), but there were also numerous lesser peaks over the range 12730-17250 Da. The peaks fell into four groups: a group in the range 12730 – 13230 Da around the major peak at 12863 Da; a group in the range 13400 – 13900 (approximately) around the major peak at 13758.7 Da; a group of lesser peaks in the range 14,000 – 14,700 Da; and a group of peaks in the range 16,800 – 17,500 Da. Unambiguous assignment of these peaks to fragments of ASAI using Findpept was not possible, and was complicated by the observation that none of the peaks appeared to be consistent with the predicted cleavage point of the α -factor pro-sequence. The “linker” region of ASAI contains a potential N-glycosylation site (N-X-S/T), and evidence presented later in this chapter shows that ASAI expressed in *P. pastoris* is glycosylated; any lectin polypeptides containing the linker region will thus have anomalous mass ion values.

The mass ion at 13487.5 Da and the major mass ion at 13758.7 Da present in the mass spectrum of ASAI could be putatively identified with the C-terminal lectin domain of ASAI (amino acid residues 153-262). These identifications are shown in table 3.2. This assignment is consistent with results obtained by expression of a construct containing only the C-terminal lectin domain (polypeptide 2; i.e. residues 153-279) of ASAI in *P. pastoris*. As shown in fig. 3.10, this construct gave a single polypeptide band on SDS-PAGE after purification, at approximately 13.5 kDa.

Proteolysis of ASAI by yeast extracellular proteinases in *P. pastoris* strain X-33 is much less specific than the cleavage of ASAII. This may reflect the polypeptide not being properly folded, or the effect of glycosylation of the linker region. Since the C-terminal lectin domain expressed alone in strain X-33 gave a product similar to ASAII

after proteolysis, it is clear that the N-terminal lectin domain and linker region in ASAI are responsible for the heterogeneity of this recombinant protein.

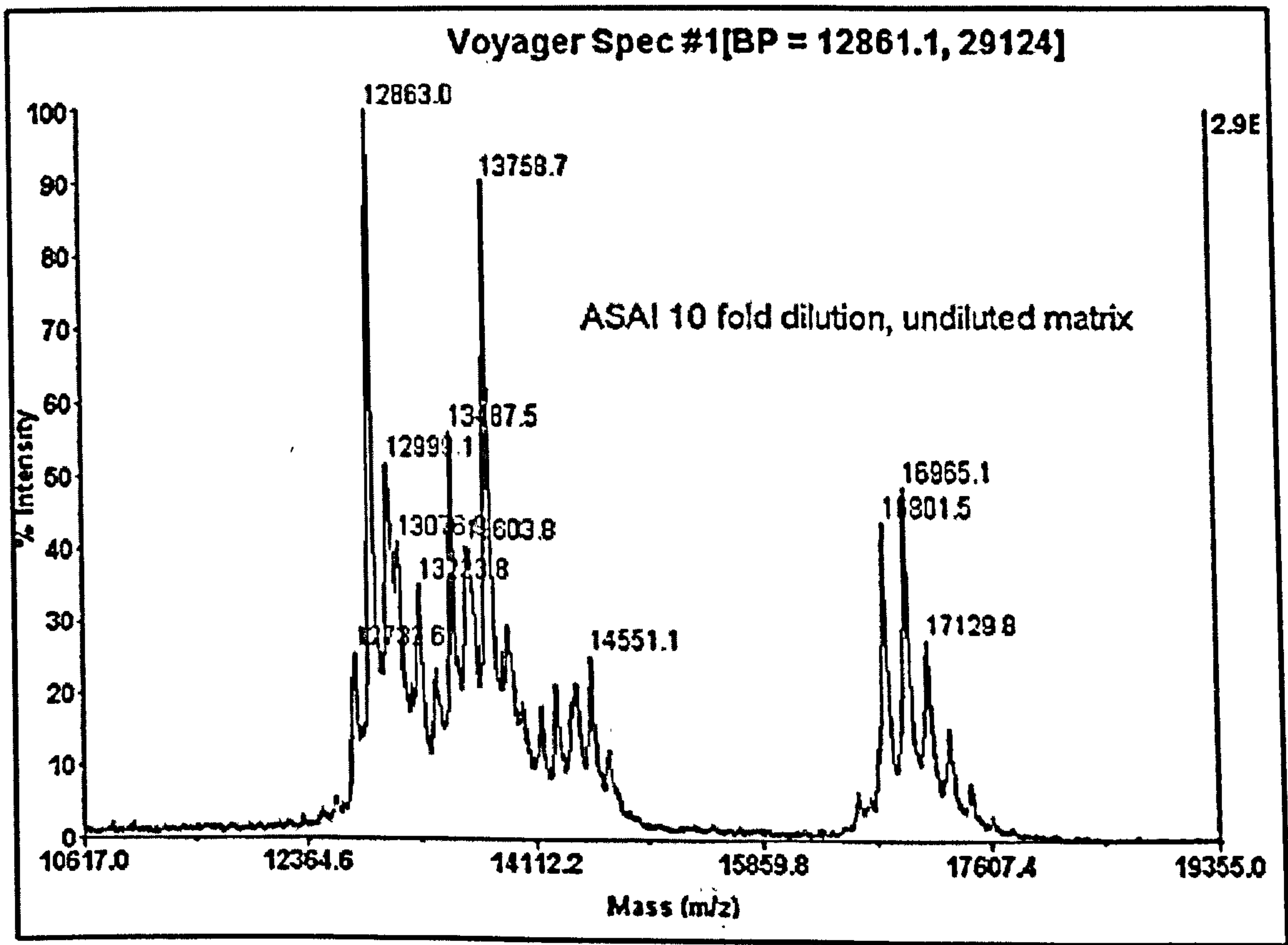


Figure 3.16 MALDI-TOF mass spectrometric analysis of recombinant ASAI (produced in *P. pastoris* strain X-33).

Expt. Mass	Predicted Mass	Residues	Peptide (Preceding and succeeding residues)
13487.5	13487.7	154-275	(R)NILRNDEGLYAGQSLDVEPYHFIMQEDCNLVLYDHSTAVWATNTDIPGKKGCRAVLQSDGNFVVYDAEGRSLWASHSVRGNGNYVLVLQEDGNIVIYGSDIWTNTYRKSVGGPVIRKVGTL(A)
13487.5	13487.7	151-271	(T)LTRNILRNDEGLYAGQSLDVEPYHFIMQEDCNLVLYDHSTAVWATNTDIPGKKGCRAVLQSDGNFVVYDAEGRSLWASHSVRGNGNYVLVLQEDGNIVIYGSDIWTNTYRKSVGGPVIRK(V)
13758.7	13757.8	154-279	(R)NILRNDEGLYAGQSLDVEPYHFIMQEDCNLVLYDHSTAVWATNTDIPGKKGCRAVLQSDGNFVVYDAEGRSLWASHSVRGNGNYVLVLQEDGNIVIYGSDIWTNTYRKSVGGPVIRKVGTLAGAA(A)

Table 3.2 Putative identification of mass ions present in MALDI-TOF mass spectrometric analysis of recombinant ASAI.

3.3.6 Glycosylation of recombinant ASAI produced in *P. pastoris* strain X-33

In order to characterise the recombinant ASAI produced in strain X-33 more closely, a gel filtration separation was carried out on a sample of protein that had only undergone partial proteolysis to low molecular weight polypeptides. Fractions from the gel filtration separation were examined by SDS-PAGE and Western blotting. Results are shown in figs. 3.17 and 3.18. The lectin eluted from the column as a broadened peak compared to ASAII, and a partial separation into lectin molecules of differing compositions occurred. The 13.5 and 14 kDa polypeptides were uniformly distributed across the broad peak, but the diffuse band of 40-65 kDa was distributed towards the earlier fractions, whereas the 32 kDa band was distributed towards the later fractions. The 16 kDa band was distributed with the 32 kDa band. Not only was the diffuse, high molecular weight band distributed towards earlier fractions, but it itself was size-fractionated. The progressive increase in polypeptide mobility across the peak fractions of the gel-filtration run shown on SDS-PAGE suggests carbohydrate molecules of decreasing size attached to the linker region of an intact ASAI proprotein; non-glycosylated ASAI proprotein is retained longer on the column and distributes towards later fractions. As the intact, glycosylated form of ASAI becomes less abundant across the gel-filtration peak with decrease in molecular weight, more intact, non-glycosylated heterodimer is present.

When selected fractions were analysed by Western blotting, and probed with anti-ASAII antibody, the bands of 13.5, 14, 32, and the diffuse 40-65 kDa region all reacted strongly, showing that they can be identified as components of ASA. These results confirm the identity of the 13.5 and 14 kDa polypeptides which are present after complete proteolytic cleavage. The 16 kDa band did not react with anti-ASAII antibodies; the reason for this is not clear, but this fragment may not contain the epitope(s) recognised by the antibody. The co-purification of this band with ASAI polypeptides through the purification steps make it unlikely to be a contaminant.

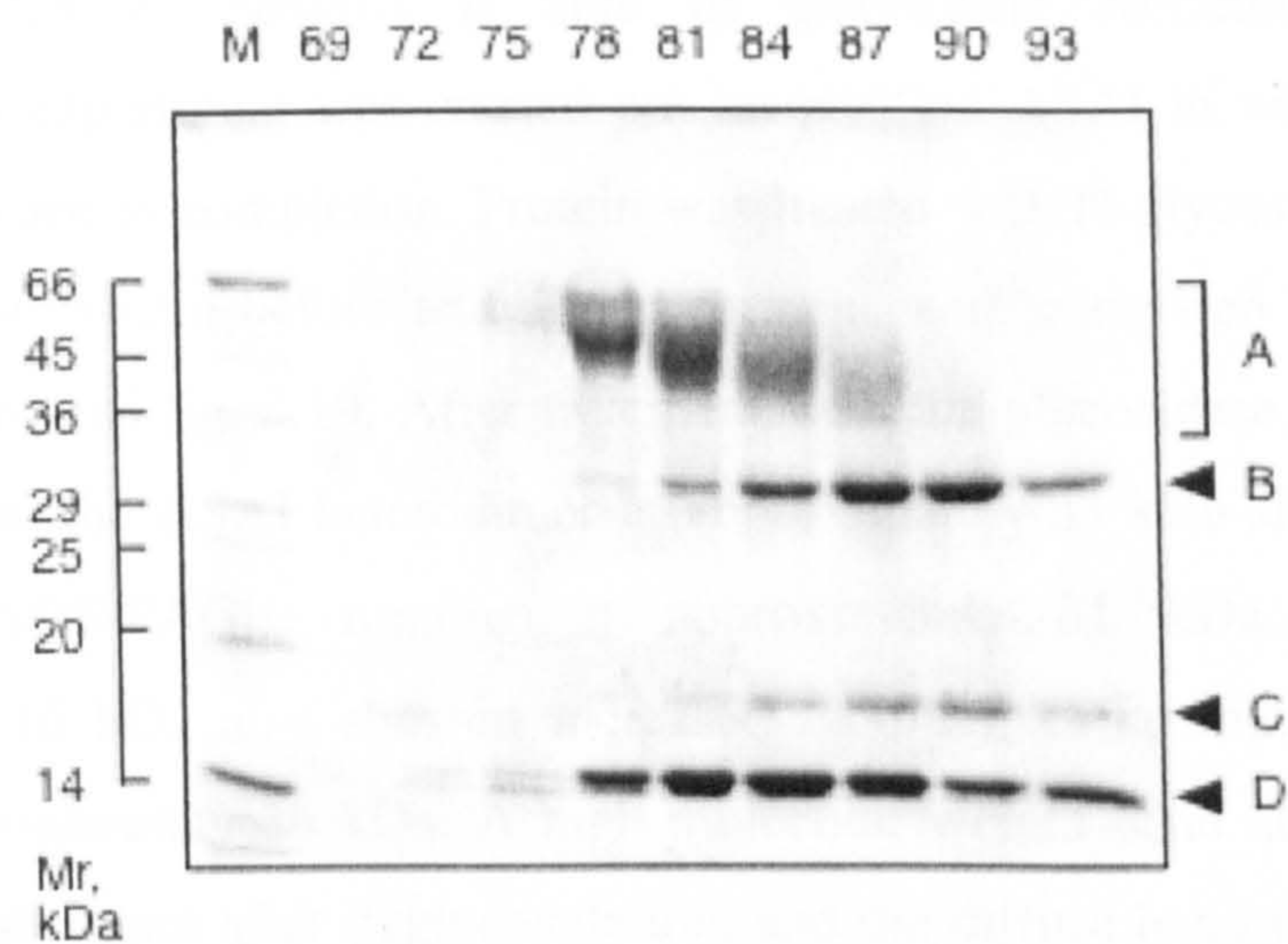


Figure 3.17 15% SDS-PAGE showing analysis of partially proteolytically cleaved ASAI produced in *P. pastoris* by gel-filtration chromatography. Numbers denote elution fraction numbers. M = SDS markers, A = glycosylated ASAI, B = Intact heterodimer, C = unidentified polypeptide, D = monomeric subunits.

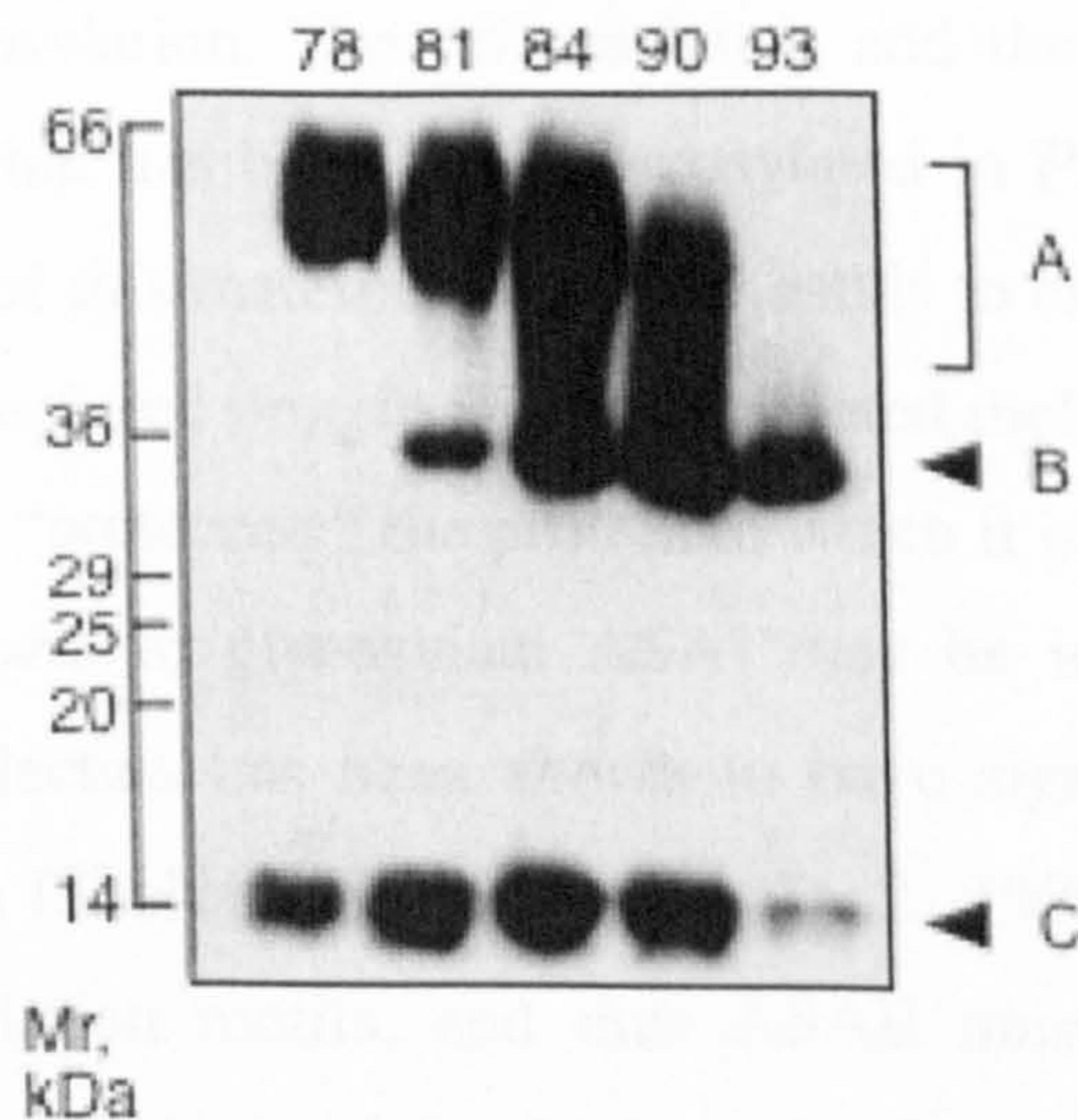


Figure 3.18 Western blot of ASAI peak fractions from fig. 3.17 above after gel-filtration chromatography, probed with anti-ASAII antibody. A = glycosylated ASAI, B = Intact heterodimer, C = monomeric subunits.

To establish that *P. pastoris* is able to glycosylate recombinant ASAI, a deglycosylation experiment was carried out on purified ASAI in which proteolysis had not taken place to completion. Protein was treated with N-glycosidase F, and the band patterns of protein before and after treatment were compared by SDS-PAGE. Results are shown in fig. 3.19. After treatment with the glycosidase, a doublet band corresponding to the ASAI heterodimer at approximately 33 kDa showed increased mobility on SDS-PAGE, running at approximately 31 kDa. The band at approximately 16 kDa also showed increased mobility, to an indicated molecular weight of approximately 14 kDa. A high molecular weight band at around 60 kDa appeared less abundant after deglycosylation, and the diffuse band in the 40-65 kDa region disappeared. The bands at 13 and 14 kDa were unaffected. These results are consistent with glycosylation of the ASAI proprotein at the N-G-T motif in the linker region (residues 123-125, fig. 3.7). The intact ASAI proprotein and the 16 kDa band contain the linker region, and their molecular weights are decreased by deglycosylation, by an amount that is consistent with “core” glycosylation in *P. pastoris*. The low molecular weight bands do not contain the linker region, and are not affected by deglycosylation. The diffuse band, and the high molecular weight band represent protein that has been hyperglycosylated in *Pichia* (Conde, R., *et al.*, 2004); deglycosylation of this material causes the bands to effectively disappear. The very heavily hyperglycosylated protein is deglycosylated inefficiently due to the large amount of carbohydrate “protecting” the protein to which it is attached.

The ability of *P. pastoris* to glycosylate ASAI may be important for activity as glycosylation of plant lectins has been shown to have significant effects on lectin structure and activation (Sheldon, P. S., Bowles, D. J., 1992). The ASAI sequence contains no N-glycosylation motifs, and thus ASAI must be a non-glycosylated protein. No evidence was observed for high molecular weight polypeptides being present on SDS-PAGE or Western blot, after nickel-affinity chromatography or gel-filtration.

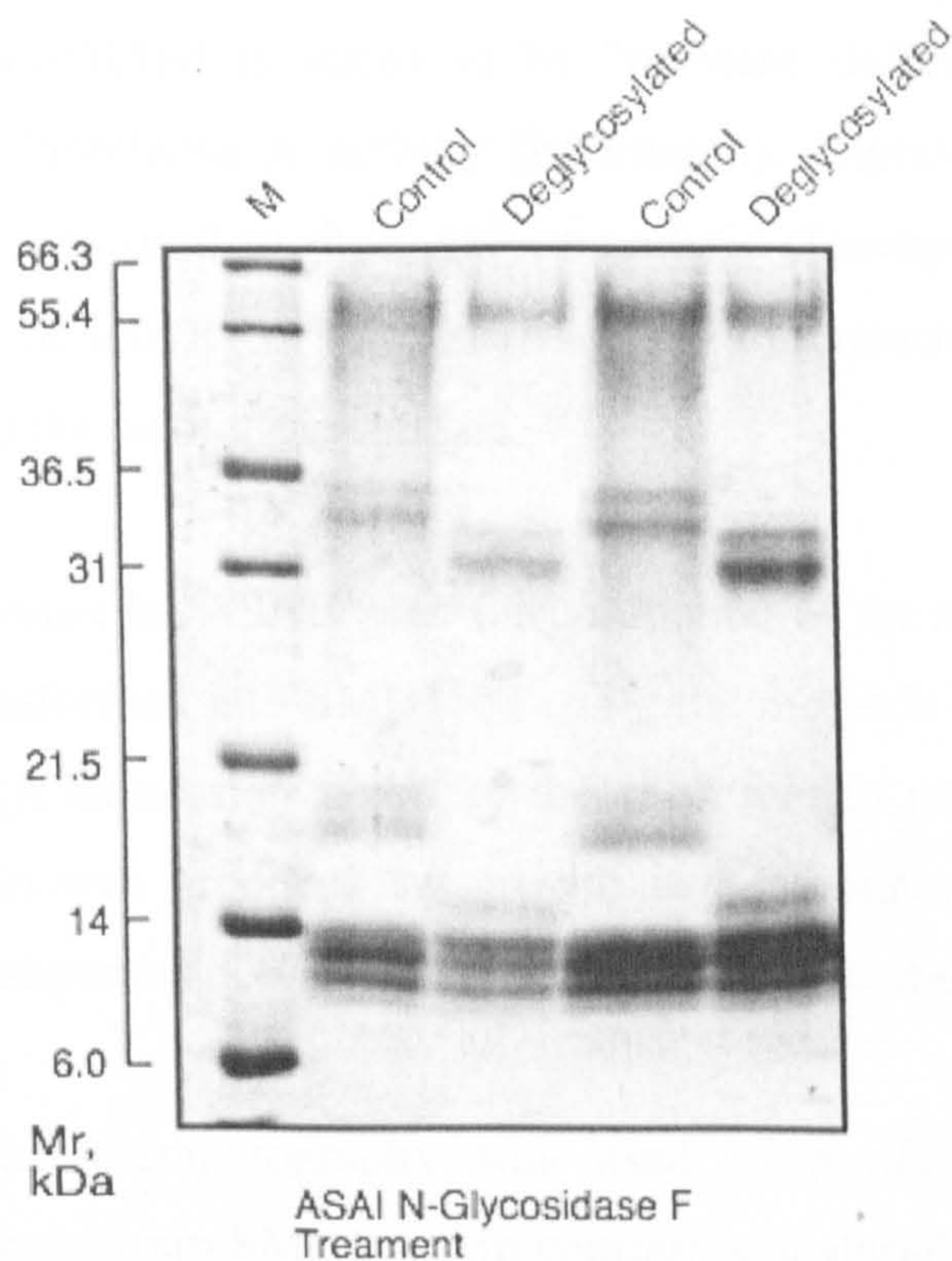


Figure 3.19 Deglycosylation of partially proteolytically cleaved ASAI produced in *P. pastoris* by treatment with N-glycosidase F; samples treated and untreated with the enzyme are compared by SDS-PAGE, followed by staining with Coomassie blue.

3.5 Purification and Characterisation of Recombinant Garlic

Lectins expressed in *P. pastoris* strain SMD1168

P. pastoris strain SMD1168 is stated to be “protease deficient”; it has a *pep4* genotype and lacks Proteinase A activity (Invitrogen). Expression of recombinant ASAII in this strain was used to show that proteolytic cleavage of the recombinant protein was a result of activity of Proteinase A, and to characterise the polypeptide initially secreted into the culture supernatant.

The expression construct for ASAII used for production of the recombinant lectins in strain X-33 was transformed into SMD1168 using the methods previously described, and clones with high expression levels of the recombinant proteins were selected. Recombinant protein was produced by growth in a laboratory scale fermentation system, and culture supernatant was taken forward for purification of proteins.

When metal affinity chromatography was used to purify recombinant ASAII produced in *P. pastoris* strain SMD1168, in contrast to material produced in strain X-33, almost all the recombinant protein bound to Ni^{2+} metal affinity columns. This was used as a method of choice for the initial purification. Typical results are shown in fig. 3.20. The protein eluted from the metal affinity column was subjected to a “clean up” step by gel filtration on Sepahacryl S-200 (fig. 3.21), as carried out for protein expressed in strain X-33.

3.4.1 SDS-PAGE of recombinant garlic lectin ASAII produced in *P. pastoris* SMD1168

Recombinant ASAII produced in *P. pastoris* strain SMD1168 was essentially a single band on SDS-PAGE, with an estimated molecular weight of 16.5 kDa, after purification by metal affinity chromatography and gel filtration. Trace amounts of polypeptides with higher and lower mol. wts. (\pm approximately 500 Da) were present. The ASAII produced in strain SMD1168 was immunoreactive with anti-myc antibodies (see fig. 3.20). The presence of the C-terminal tag region, and the molecular weight estimate in line with that predicted for ASAII which has only been

processed by removal of the α -factor pro-sequence (predicted 17140 Da) is consistent with absence of Proteinase A activity in the *pep4* yeast strain.

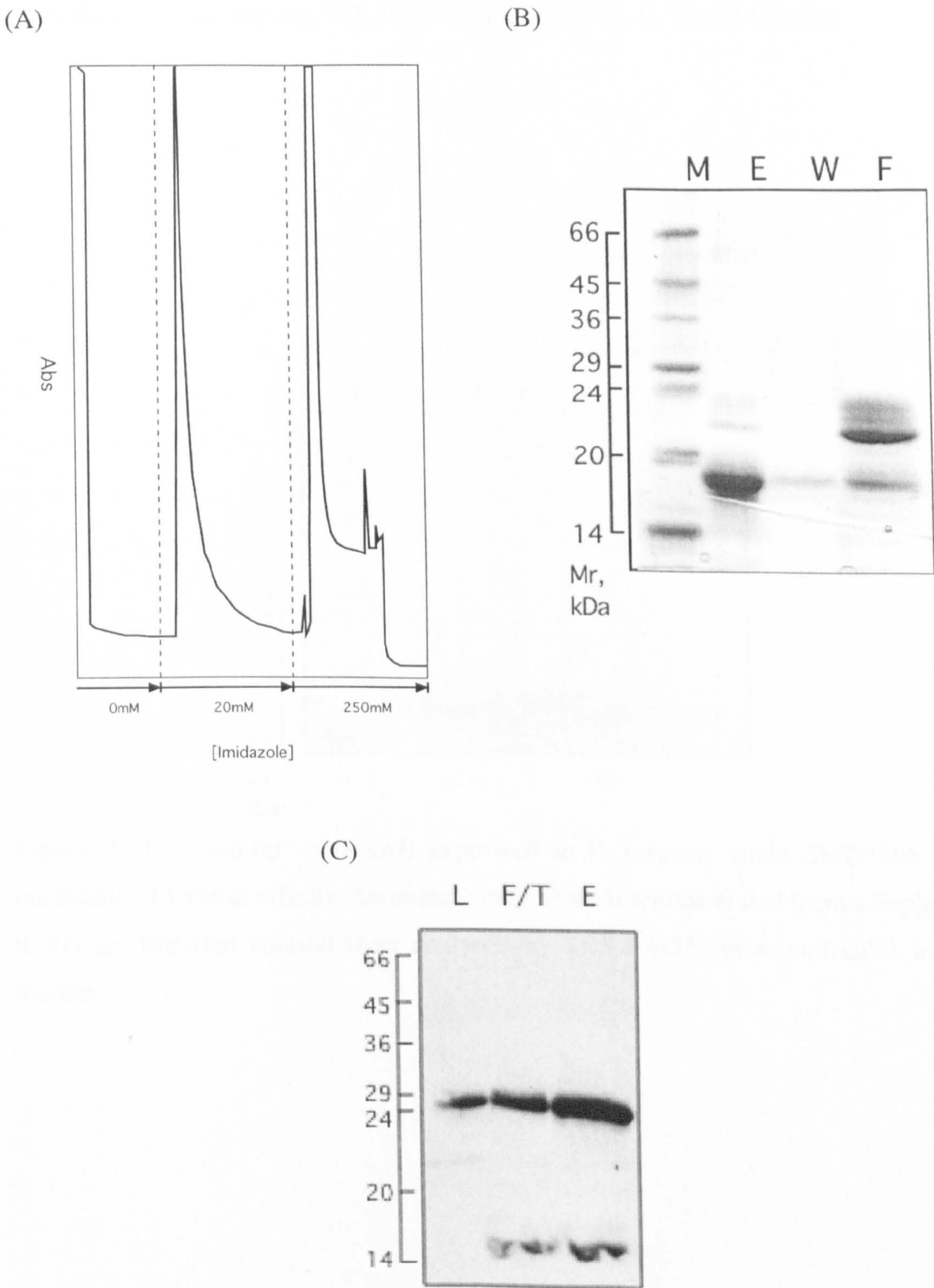


Figure 3.20 Purification of recombinant ASAII expressed in *P. pastoris* strain SMD1168. (A) Elution profile of Ni^{2+} affinity column. (B) SDS-PAGE analysis of

fractions eluted from column: F, flow-through fraction (non-bound); W, wash fraction (20mM imidazole); E, protein eluted at 250mM imidazole; M, molecular weight marker. (C) Western blot analysis of Ni² affinity purification using anti-myc epitope antibodies: L, Load fraction, F/T, flow-through fraction, E, Eluted fraction.

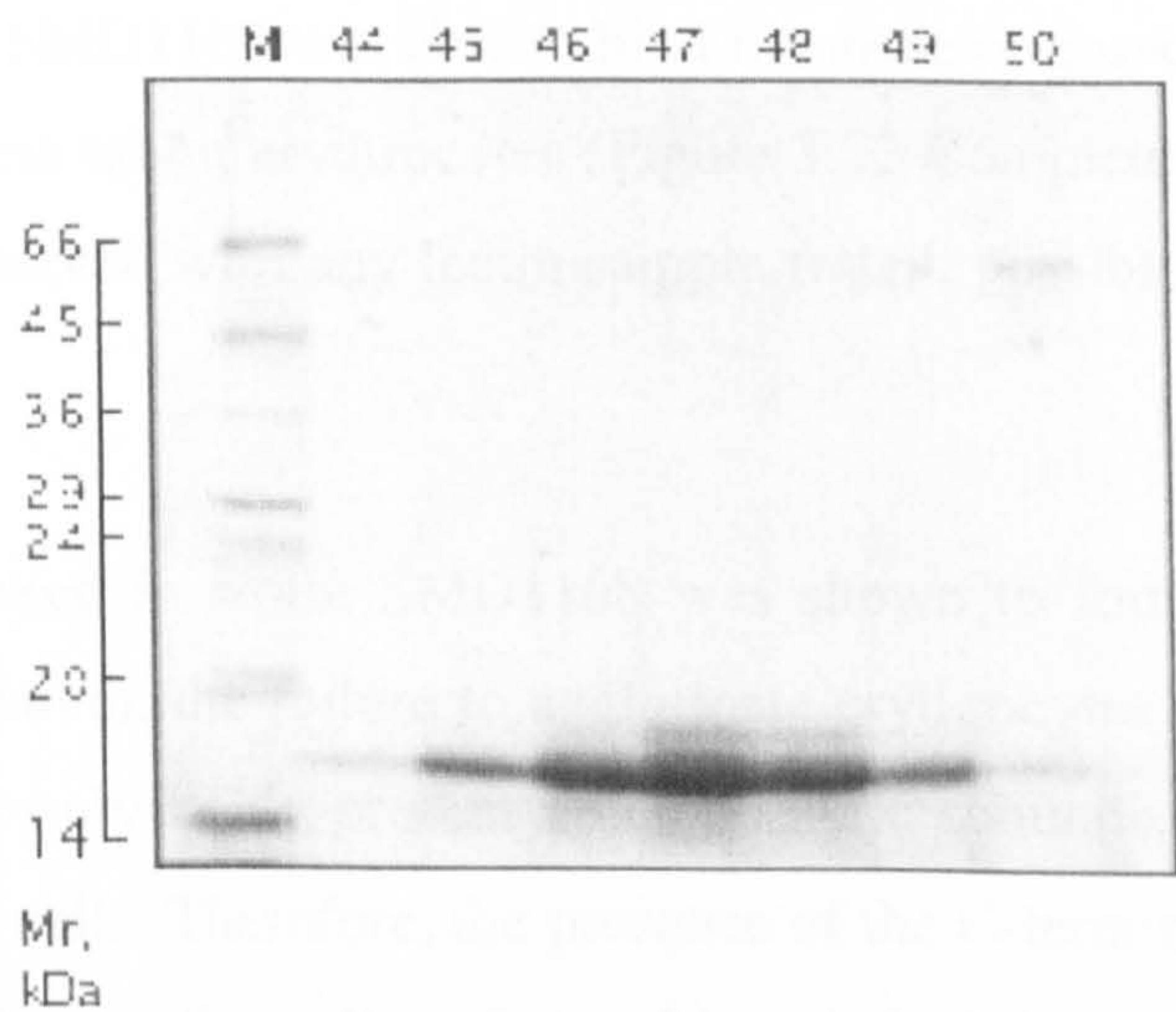


Figure 3.21 “Clean-up” of ASAIL expressed in *P. pastoris* strain SMD1168 after purification by metal affinity chromatography. Peak fractions eluted from a Sephacryl S-200 gel filtration column were analysed by SDS-PAGE. M = molecular weight marker.

3.4.2 Functional properties of recombinant garlic lectin ASAII produced in *P. pastoris* strain SMD1168

ASAII expressed in SMD1168 could not agglutinate rabbit erythrocytes at a concentration of 2.5 mg/ml (result not shown). However, the protein was functionally active, since it was able to bind to mannan-agarose.

A pull-down experiment using mannan-agarose was used to demonstrate binding to a mannose oligosaccharide. 125 µg of recombinant lectin was incubated with mannan-agarose in 100 µl PBS and after overnight incubation the sample was spun down and the supernatant and pellet run on SDS-PAGE. This demonstrated that recombinant ASAII expressed in SMD1168 was able to bind mannose oligosaccharides, despite its inability to agglutinate rabbit erythrocytes (Figure 3.22 Complete binding to mannan-agarose was not observed with any lectin sample tested, possibly due to overloading the matrix).

Since ASAII expressed in strain SMD1168 was shown to form dimeric molecules (see section 3.2.3 above), the failure to agglutinate erythrocytes cannot be due to the recombinant ASAII only being present as monomeric subunits, and therefore being unable to cross-link cells. Therefore, the presence of the C-terminal “tag” region must interfere with binding to the cell surface, although it does not affect the protein’s ability to bind to mannose oligomers.

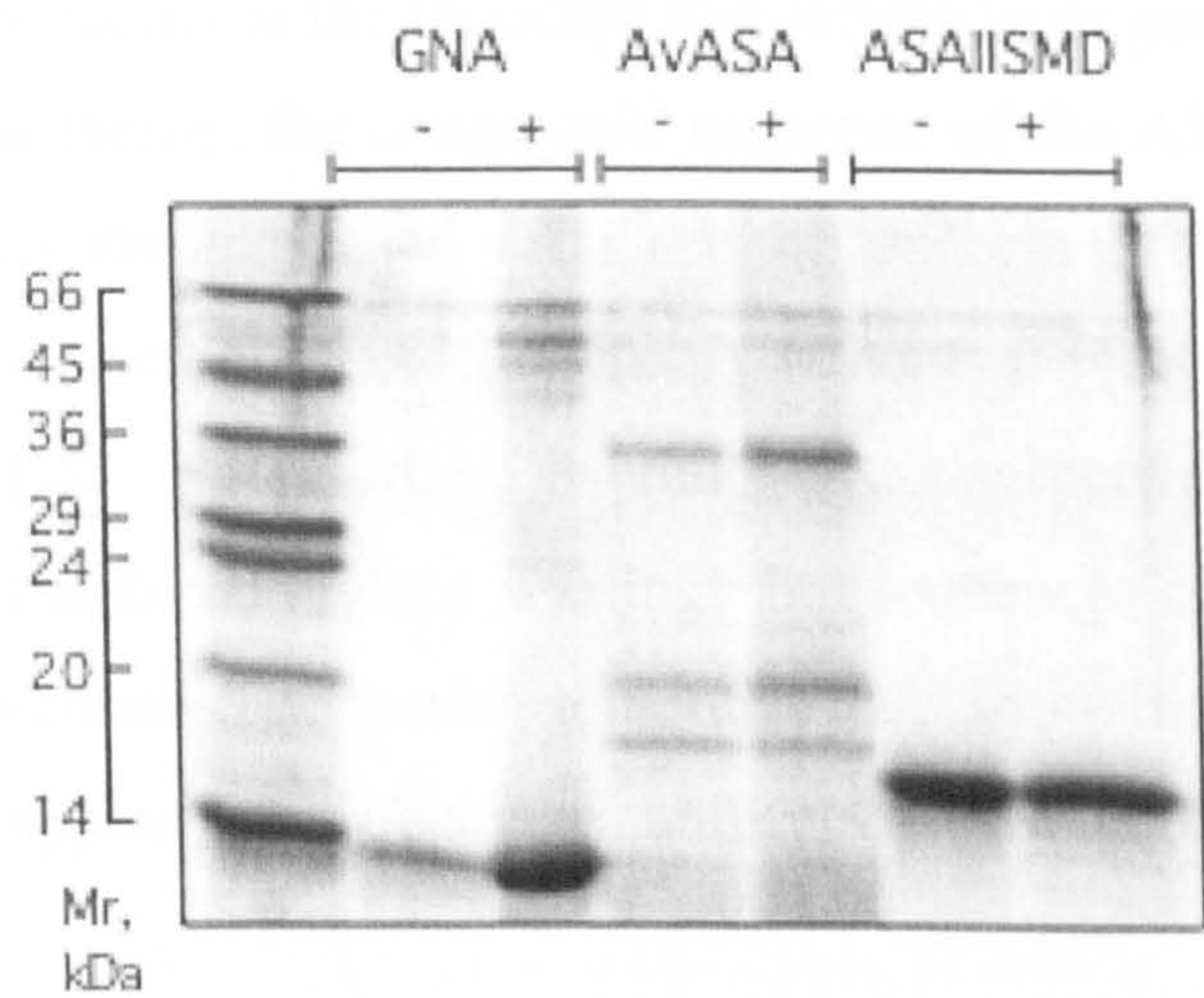


Figure 3.22 SDS-PAGE showing mannan-agarose binding of recombinant garlic lectins expressed in *P. pastoris* strain SMD1168. - = non-bound protein, + = bound protein: GNA, *Galanthus nivalis* agglutinin; AvASA, avidin-ASA fusion protein; ASAIISMD, ASAII expressed in *P. pastoris* strain 1168.

3.5 Discussion

The results of amplification of lectin genes from garlic bulb genomic DNA presented here support previous claims in the literature that lectins from garlic bulbs are part of a multi-gene isoelectin family, the amino acid sequence of the ASAI gene containing 29 substitutions from the published sequence of isoform 3 (M85176.1). Similar isoelectin variation has been observed previously (Chandra, N. R., *et al.*, 1999), (van Damme, E. J. M., *et al.*, 1992), where overall homology between lectin nucleotide sequences and 3' untranslated regions was high but sequences contained variation in some amino acid positions. The results presented here show that the 29 amino acid variations do not appear to affect the ability of ASAI to agglutinate rabbit erythrocytes.

Despite the large amounts of mannose-polysaccharides produced by *P. pastoris* production of recombinant garlic lectins by laboratory scale fermentation was a viable strategy for production of these proteins at the milligram level. This is a lower yield than has been reported for some other plant lectins such as the kidney bean (*Phaseolus vulgaris*) agglutinin E-form (PHA-E), which has been expressed at approximately 100 mg/L (Baumgartner, P., *et al.*, 2002) and GNA has also been expressed at high levels in *P. pastoris*, approximately 80 mg/L (Baumgartner, P., *et al.*, 2003). Further optimization of recombinant garlic lectin expression could have been achieved by a large-scale screen of transformants but with yields of between 10-20 mg/L of culture supernatant, the amounts of recombinant protein produced were amenable to lab-scale purification and sufficient for insect feeding trial experiments.

Variable levels of post-translational processing of recombinant garlic lectins were observed in the X33 strain of *P. pastoris*. Complete processing of ASAI *in planta* produces two lectin polypeptides of 11.5 and 12.5 kDa. In some instances of ASAI production this was observed (fig 3.12), whereas in others incomplete processing was observed (fig 3.17). The work of Bacchawat (Bacchawat, K., *et al.*, 2001) showed that ASAI dimerization was reversible and that unfolding into monomeric subunits essentially occurs via a simple two-step process from folded dimer to unfolded monomeric subunits, no intermediate states were observed suggesting that the energy

of stabilization in the dimer derives directly from the monomer-monomer interface. The hydrophobic core of the ASAI dimer contains three conserved tryptophan residues which are brought together via strand exchange of the monomeric subunits. ASAI dimers possess three 4-stranded anti-parallel β -sheets which form three sides of a 12-stranded β -barrel. Strand exchange occurs between three strands of the first sheet which interact with a strand which crosses from the adjacent sheet. This process has also been termed 'domain swapping' (Bennett, M. J., *et al.*, 1994) and has a major consequence for stabilizing the dimer. The observation of variation in the denaturation of ASAI on SDS-PAGE may be accounted for by partial breaking of the dimeric interface at the hydrophobic core, as once this association is broken down ASAI monomers do not retain appreciable secondary or tertiary structure (Bachhawat, K., *et al.*, 2001). The ambiguous MALDI-TOF data obtained for ASAI X33 may be attributable to proteolysis resulting from this incorrect folding of the polypeptide. ASAI was also correctly glycosylated in *P. pastoris*, the exact role of glycosylation of ASAI is not known, some plant lectins are inactivated by glycosylation, for example non-glycosylated pro-Con A is known to be active but glycosylated pro-Con A is inactive suggesting that glycosylation alone is sufficient to inactivate Con A carbohydrate binding. ASAI produced in *P. pastoris* was able to agglutinate rabbit erythrocytes, however a heterogeneous mixture of glycosylated and non-glycosylated lectin was used and so it remains unclear whether ASAI is inactivated by glycosylation or not. Correct processing of ASAII *in planta* produces a single 12 kDa polypeptide. In the X33 strain this processing appeared to take place correctly.

C-terminal post-translational processing of both recombinant lectins took place in the X33 strain of *P. pastoris*. C-terminal proteolysis is a common modification of plant lectins *in planta* and has been observed in the case of GNA where a 105 amino acid mature polypeptide was produced from cDNA (van Damme, E. J., *et al.*, 1991) and a 109 amino acid polypeptide identified using a crystallographic approach (Hester, G., *et al.*, 1995) suggesting cleavage had occurred during expression of GNA105, however, agglutination activity was comparable between the two different forms when both isoforms were expressed in *P. pastoris* and compared (Raemaekers, R. J., *et al.*, 1999). A C-terminal peptide is also cleaved off as part of the complex post-translational modifications carried out on Con A *in planta*. Interestingly, C-terminal

cleavage in recombinant garlic lectins occurred between two residues N-terminal of the site observed *in planta* between two hydrophobic residues Val 118 and Ile 119. Proteinase A is a pepsin-like aspartic protease present which is secreted into the extracellular medium in yeasts with broad specificity due to its extended binding pocket. The encoding gene *pep4* is present in the X33 strain of *P. pastoris*. Proteolysis of the lectin C-terminus *in planta* occurs between two basic residues, Arg 120 and Lys 121 rather than the hydrophobic residues in yeast implying that proteolysis may be carried out in accordance with the structure of the protein rather than particular enzyme specificity. This would suggest that the C-terminus of garlic lectin is not integral to the folded polypeptide but rather occurs near the exposed surface of the protein and is susceptible to proteolysis by a range of proteolytic enzymes. The absence of the C-terminus in garlic lectins does not affect lectin activity, suggesting that it may have an initial role in protein stability or folding but is not required for carbohydrate binding but the absence of agglutination activity in ASAII SMD1168 suggests that the presence of the C-terminus may interfere with lectin activity although this may be accounted for by the bulky C-terminal tag extension in ASAII SMD1168. The greater reproducibility of ASAII expression makes it more suitable for further use in recombinant fusion proteins. The homogeneity of the ASAII expression product makes analysis of any potential insecticidal effects and protein-protein interactions with the insect gut easier to interpret, and the information obtained about the nature of C-terminal cleavage of lectins will be valuable for preventing the proteolysis of recombinant fusion proteins in the insect gut. Variable post-translational processing of ASAI makes it inconvenient for use in insecticidal fusion proteins even though an intact heterodimeric garlic lectin may be a better carrier molecule for insecticidal proteins or peptides, problems with cleavage of ASAI would make analysis of the effects of fusion proteins ambiguous.

Chapter 4

Biological Activity of Recombinant Garlic Lectins on Insects

4.1 Introduction

There has been considerable interest in the past in the insecticidal action of plant proteins towards insect species. Not only is this relevant to understanding how plant-insect interactions are determined in Nature, but attempts have also been made to exploit plant proteins as insecticides, primarily through expression in transgenic plants of a different species to that in which the protein is normally produced. Exploitation of plant insecticidal proteins has largely been overtaken by the widespread use of the protein toxin from the bacterial species *Bacillus thuringiensis* (*Bt*), which has been demonstrated to be toxic to lepidopteran (Nayak, P. *et al.* 1997) and coleopteran insects, both when applied as a topical insecticide, and when expressed in transgenic plants (Tu, J. *et al.* 2000). However, the spectrum of insecticidal activity shown by *Bt* toxins is limited, and no *Bt* toxin has been reported to show toxicity towards hemipteran insects. *Bt* toxin requires activation by proteolysis of a protoxin, and it has been suggested that the low level of protease activity in the guts of hemipteran insects does not allow this activation to occur, so that subsequently no toxicity is observed. The absence of a usable *Bt* toxin to protect crops against hemipteran pests has led to continuing study of the toxic effects of plant lectins towards insects, as an alternative strategy for insect control.

The mannose-binding lectin family, exemplified by snowdrop lectin (*Galanthus nivalis* agglutinin; GNA) have been shown to have significant insecticidal activity against hemipteran insects, both when delivered via artificial diet (Powell, K. S. *et al.* 1993), and via expression in transgenic plants (Rao, K. V. *et al.* 1998). In understanding the effects of lectins on insect pests it is important to evaluate a specific lectin against a specific pest. GNA has only marginal effects on lepidopteran pests (Fitches, E. C. *et al.* 1997), whereas it significantly decreased survival in most hemipterans tested. However, even between species from the same order, differences in toxicity are apparent; at concentrations which would induce high levels of mortality in rice brown plant hopper, GNA has only marginal effects on the survival of the

potato aphid *Aulacorthum solani* (Down *et al.* 1996). Similarly, despite high levels of sequence similarity between members of the mannose-binding lectin family, it is interesting to note that the level of toxicity of different lectins towards a single insect species may vary significantly; for example, garlic lectin is at least 10-fold less toxic (on a concentration basis) towards rice brown plant hopper than GNA (Powell *et al.* 1995).

In this chapter the effects of recombinant garlic lectins ASAI and ASAII are tested against pea aphid (*Acyrtosiphon pisum*), peach-potato aphid (*Myzus persicae*) and rice brown plant hopper (*Nilaparvata lugens*) in artificial diet feeding bioassays. Further *in vivo* analysis was carried out by identification of ASAII binding proteins in the gut of *A. pisum*, and transport of ASAII through the gut lumen into the aphid haemolymph is examined using an avidin-ASAII fusion protein.

4.2 Insect Artificial Diet Bioassays

4.2.1 Artificial Diet Bioassays Against Symbiotic and Aposymbiotic Pea Aphid (*Acyrtosiphon pisum*)

Bioassays carried out to assess the toxicity of garlic lectins towards aphids have given variable results, but a consistent negative effect on survival and development has been observed. In addition, differences in toxicity between the different garlic lectin isoforms has been claimed, with garlic leaf lectin (ASAL) being consistently more toxic than other isoforms. It has been suggested by Banerjee and co-workers (Banerjee, S., *et al.*, 2004) that the observed toxicity of ASAL towards mustard aphid (*Lipaphis erysimi*) was based on an interaction of the lectin with a chaperonin GroEL-like protein produced by the obligate endosymbiont *Buchnera aphidicola*, called symbionin.

The pea aphid, *Acyrtosiphon pisum*, also harbours *B. aphidicola*, and thus was selected as a suitable system to test the insecticidal activity of the recombinant garlic lectins produced in *Pichia pastoris*, and to investigate whether the presence of symbiotic bacteria affected lectin toxicity. Artificial diet bioassays were carried out on neonate nymphs (0-24 h) of both symbiotic and aposymbiotic *A. pisum* to reveal

whether or not a differential response could be seen between aphids containing symbionin and aphids from which symbionin was absent. Aphids were rendered aposymbiotic by including a suitable antibiotic in the diet; garlic lectins used in the assay were produced by expression in *P. pastoris* strain X-33, and purified by ion exchange chromatography and gel filtration. Results are summarized in figures 4.1 – 4.4.

In artificial diet feeding bioassays carried out on neonate nymphs showed that standard diet could support aphid growth and development giving survival for over 17 days. Day 0 neonate nymphs possess minimal nutrients from the plant host. Any effects of lectins on these aphids will be most apparent. Neonate aphids survived for 2-3 days when diet was not available, but dehydration was prevented, giving a basis for mortality by starvation. Bioassays were carried out using ASAI and ASAII at the following concentrations: 2 mg/ml, 0.5 mg/ml and 0.125 mg/ml.

In bioassays against symbiotic insects, ASAI and ASAII showed a significant effect upon survival ($p > 0.05$, survival analysis) on *A. pisum* at all concentrations tested. Survival on garlic lectin-containing diets was reduced in comparison to control insects in a dose-dependent manner; survival was most reduced at 2 mg/ml, whereas survival plots at 0.5 mg/ml and 0.125 mg/ml were either similar, or survival at the lower dose was better. The effects of garlic lectins on aphids can be divided into three phases; acute toxicity, up to approximately 3-4 days exposure, and two phases of chronic toxicity, the first up to approximately 10 days, corresponding to the time taken for the neonate aphids to complete development, the second over the remainder of the assay (up to 20 days), during which time the aphids produced young. Survival in control aphids (i.e. diet with no added lectin) was >90% over the acute phase and the first chronic phase, but declined gradually over the period of nymph production to approximately 50% after 17 days. ASAII at the highest dose, 2.0 mg/ml, showed significant toxicity during the acute phase of exposure, with survival falling to 55% after 2 days. This was the only treatment to show acute toxicity to symbiotic pea aphids. Over the second phase, all treatments showed a negative effect on survival, which was dose-dependent for each lectin. ASAII was more effective than ASAI, with survival at 9 days for ASAII being 0%, 45%, 75%, and 90% for 2.0 mg/ml, 0.5 mg/ml, 0.125 mg/ml and 0 mg/ml (control) respectively, whereas for ASAI survivals

were 30%, 60%, 100% and 100% for 2.0 mg/ml, 0.5 mg/ml, 0.125 mg/ml and 0 mg/ml (control) respectively. All ASA treatments showed a marked effect on aphid survival during the third phase of the assay, with survival falling much faster than in controls, and all aphids exposed to ASA being dead after 17 days.

When aposymbiotic *A. pisum* were assayed under the same conditions, survival of the controls was 100% during the acute phase of the assay (to 5 days) but declined significantly subsequently, with all aphids being dead by day 20. The survival of aposymbiotic controls on diet is significantly different from the survival of symbiotic controls ($p < 0.05$, survival analysis). Poorer survival of the aposymbiotic aphids would be expected in the assay, as the absence of symbionts interferes with normal nitrogen metabolism. *B. aphidicola*, the obligate endosymbiont, is known to be involved in the assimilation of certain amino acids in the aphid and its absence causes a decrease in survival (Douglas, A. E., *et al.*, 2001); the amino acid mix in the diet is unable to entirely compensate for the absence of nutrients normally supplied by the symbiont. Nevertheless, differences in the survival curves between control aposymbiotic aphids and aposymbiotic aphids exposed to garlic lectin were still apparent and significant ($p < 0.05$, survival analysis), except for the lowest concentration of ASAI tested. In all cases aphids exposed to garlic lectin survived less well than controls. The acute toxicity of the highest dose of ASAI, 2.0 mg/ml, was observed in assays with aposymbiotic aphids, with survival reduced to 55% after 2 days and 40% after 4 days; lower doses of ASAI showed their major effects on survival during the chronic toxicity phases, with survivals at 9 days being 10%, 25%, 35% and 45% for ASAI doses of 2.0 mg/ml, 0.5 mg/ml, 0.125 mg/ml and 0 mg/ml (control) respectively. All ASAI treatments resulted in total mortality of aphids by day 13, at which point 25% of control aphids survived. ASAI showed limited acute toxicity to aposymbiotic aphids (80% survival after 2 days), and chronic toxicity was shown by ASAI at 0.5 mg/ml and 0.125 mg/ml, survival after 9 days being 35% for both doses, compared to 45% for the 0.125 mg/ml dose, and 55% for controls. As was the case for ASAI, all ASAI treatments resulted in total mortality of aphids by day 13, at which point 20% of control aphids survived.

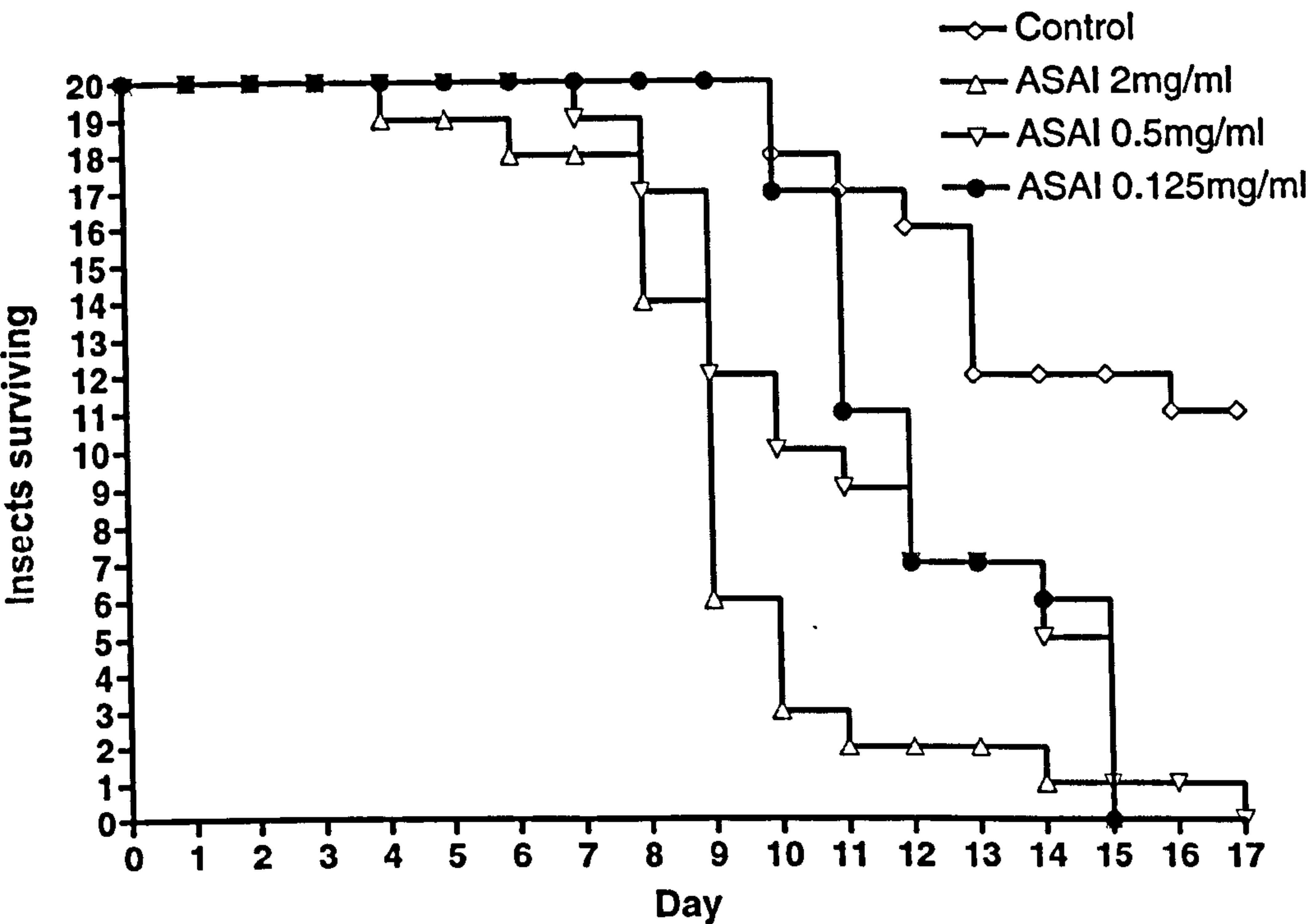


Figure 4.1 Artificial diet bioassay of recombinant ASAI against symbiotic *A. pisum*.

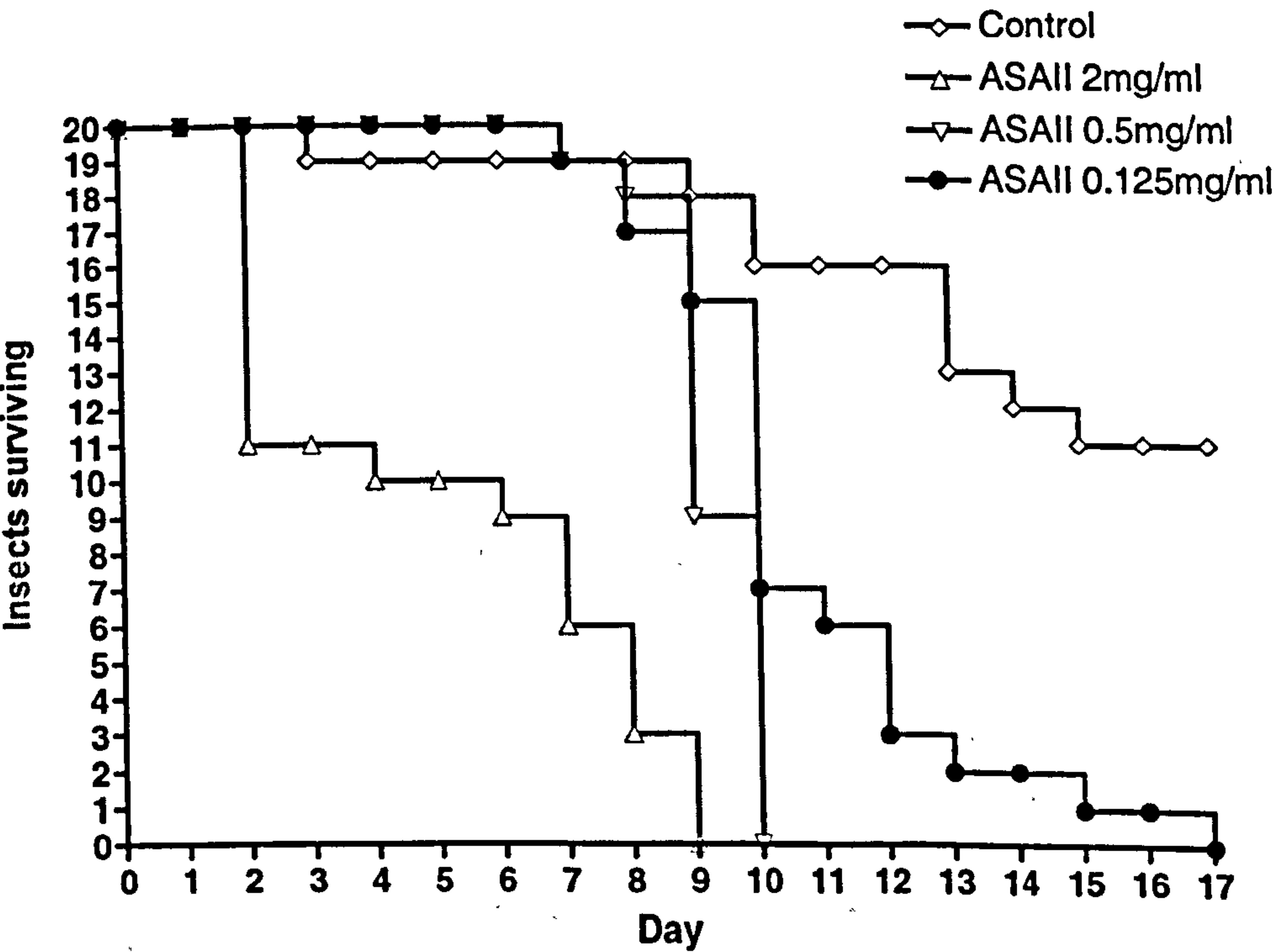


Figure 4.2 Artificial diet bioassay of recombinant ASAI against symbiotic *A. pisum*.

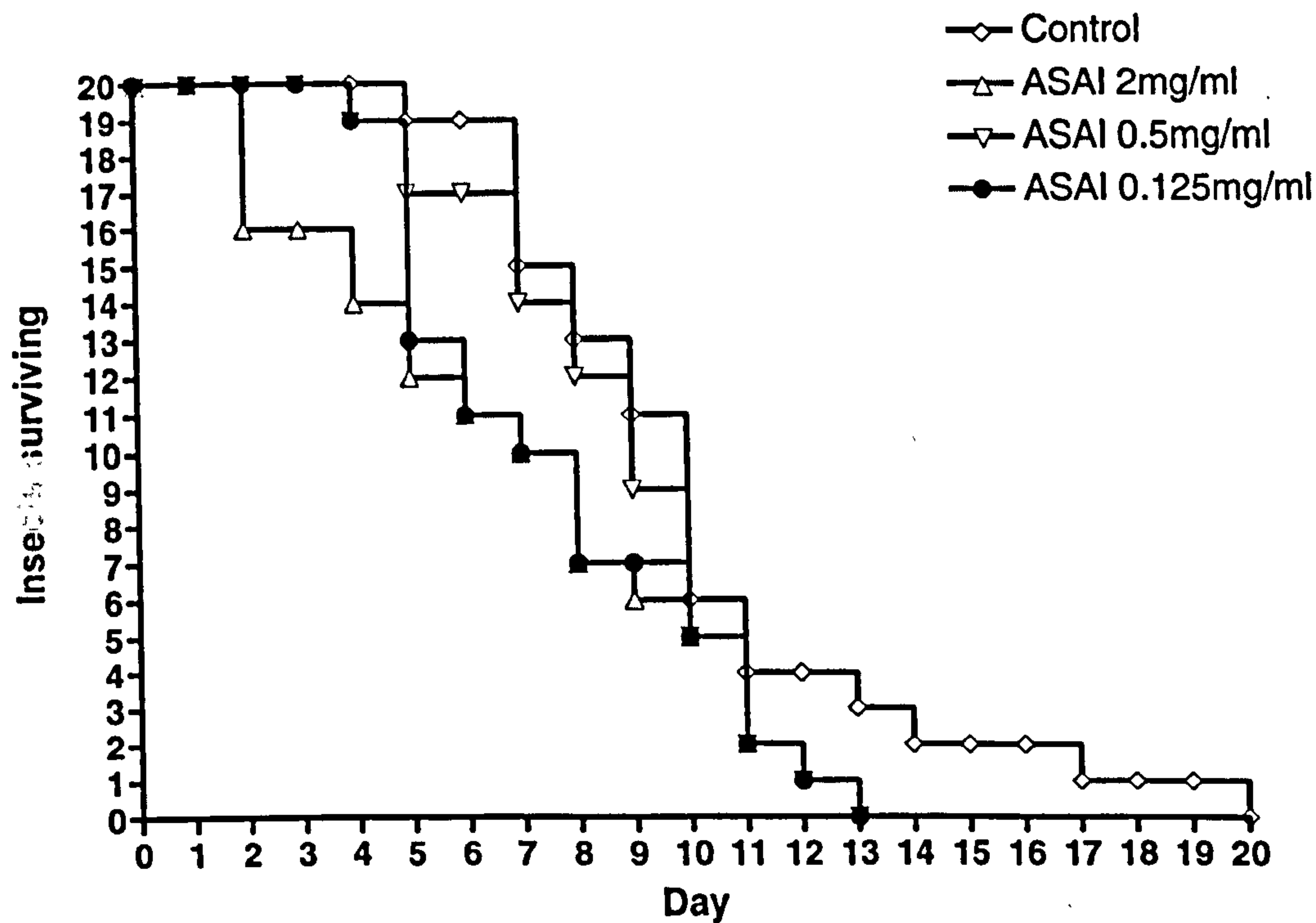


Figure 4.3 Artificial diet bioassay of recombinant ASAI against aposymbiotic *A. pisum*.

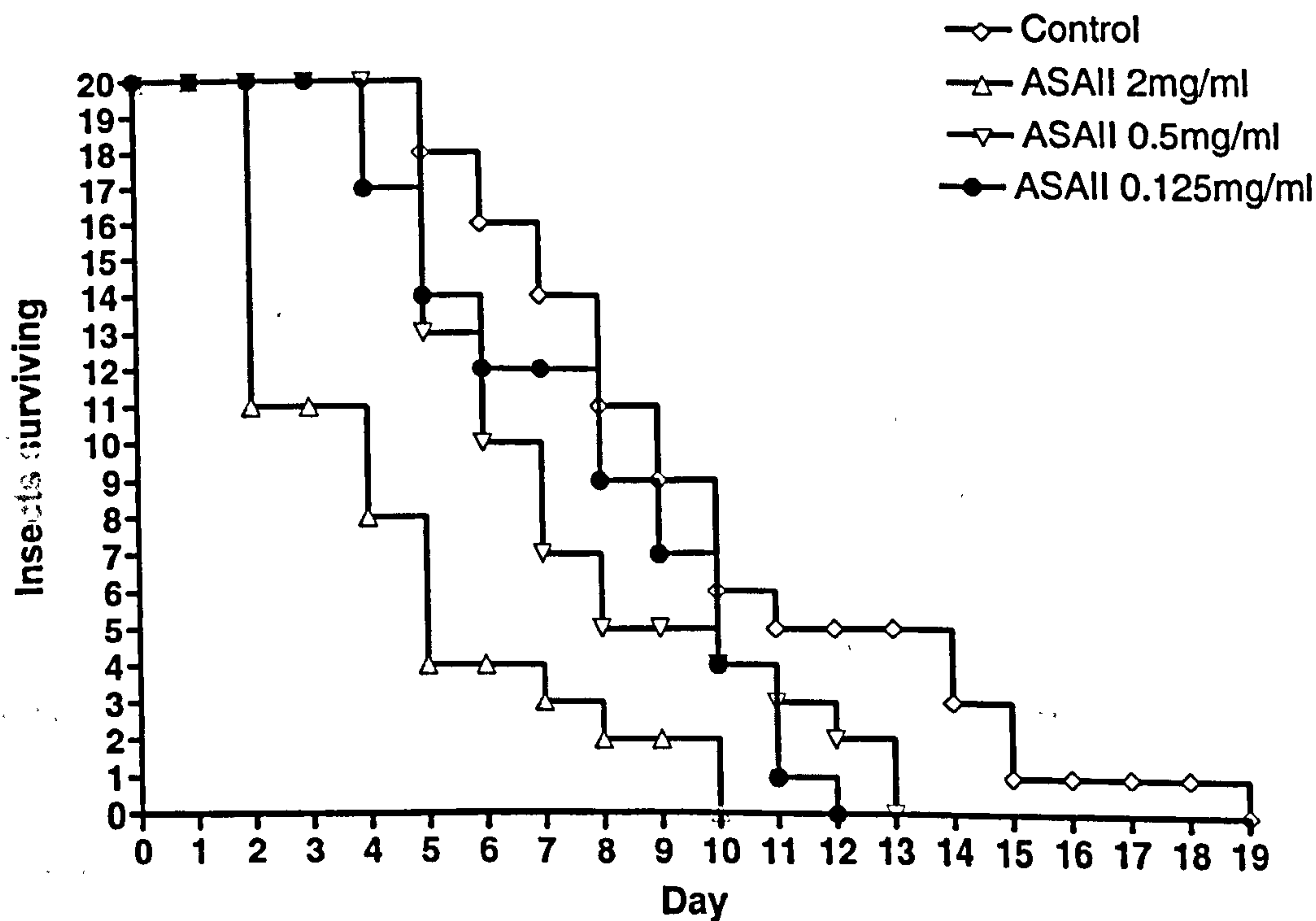


Figure 4.4 Artificial diet bioassay of recombinant ASAI against aposymbiotic *A. pisum*.

The results of the bioassays of recombinant garlic lectins against symbiotic and aposymbiotic *A. pisum* show that both ASAI and ASAI lectins are significantly insecticidal against this insect, in a dose-dependent manner, although ASAI is significantly more insecticidal than ASAI when survival curves are compared for identical dosages. Comparison of results between symbiotic and aposymbiotic aphids is complicated by the difference in survival of aphids not exposed to lectins, but it is apparent that the lectins have significant, dose-dependent acute and chronic toxicity towards the aposymbiotic aphids, which is similar to that shown towards symbiotic aphids. There is no evidence to suggest that the insecticidal effects of garlic lectins on aposymbiotic aphids are significantly less than their effects on symbiotic aphids.

4.2.2 Artificial diet bioassays against symbiotic and aposymbiotic peach-potato aphid (*Myzus persicae*)

Artificial diet bioassays were carried out on symbiotic and aposymbiotic *Myzus persicae* in a similar way to *A. pisum*. Neonate nymphs (0-24 h) were placed on artificial diet with or without the addition of rifampicin. Concentrations of recombinant lectin tested were the same as for *A. pisum*, 2 mg/ml, 0.5 mg/ml and 0.125 mg/ml. *M. persicae* bioassays were carried out at the University of York by Prof. Angela Douglas, where each treatment was tested on 10 insects. Results are summarized in figures 4.5 - 4.7.

M. persicae develops more quickly than *A. pisum*, and the phases of the assay were slightly different. The acute phase was of similar duration, up to approximately 4 days, with a similar survival under "starvation" for both aphids. Development to adult in *M. persicae* takes approximately 8 days, so the first "chronic" phase is of shorter duration. Survival times for symbiotic adult *M. persicae* on control diets extended to 50 days, although a standard "cut-off" of 21 days is used for the data presented. As was the case for symbiotic and aposymbiotic *A. pisum*, survival of aposymbiotic *M. persicae* was significantly poorer than for symbiotic aphids, and in one assay (against ASAI) aposymbiotic *M. persicae* showed very poor survival, with all aphids dying after 9 days. Data from this bioassay are not presented.

In assays with symbiotic aphids, survival on control diet was $\geq 90\%$ over 14 days and $\geq 60\%$ over 21 days. Both lectins showed limited acute toxicity ($\leq 20\%$ mortality), but had a major effect on survival during the second chronic phase of the assay, when aphids were producing nymphs. All aphids on the highest concentrations of lectins tested (2.0 mg/ml) were dead by day 8 (ASAI) or day 9 (ASAI). ASAI had little effect at the lower concentrations tested until extensive nymph production had taken place, when a sudden drop in survival took place, with all aphids dead at day 21. ASAI showed a progressive decrease in survival after day 8, with all aphids exposed to 0.5 mg/ml ASAI dead by day 12, and all aphids exposed to 0.125 mg/ml ASAI dead by day 15. ASAI was consistently more insecticidal than ASAI, as observed for *A. pisum*. All survival curves were significantly different to the appropriate control ($p < 0.05$; survival analysis).

In the bioassay of aposymbiotic *M. persicae* against ASAI control survival showed a steady decline over the assay, but did not reach zero until day 26; comparisons with lectin treatments were thus possible. The survival curve for insects on lectin-containing diets at 0.5 mg/ml and 2.0 mg/ml were significantly different to the control, whereas at the lowest lectin concentration tested, 0.125 mg/ml, no significant difference was observed. Once again, acute toxicity of ASAI towards this aphid was not significant, but the major effects on survival were seen at the onset of nymph production. All aphids exposed to ASAI at 2.0 mg/ml were dead by day 9, and all aphids exposed to ASAI at 0.5 mg/ml were dead by day 12.

The results of the bioassays of recombinant garlic lectins against symbiotic *M. persicae* are similar to those for *A. pisum*. Both ASAI and ASAI lectins are significantly insecticidal, in a dose-dependent manner, and ASAI is significantly more insecticidal than ASAI. However, *M. persicae* is less sensitive to the acute toxic effects of garlic lectin than *A. pisum*. Comparison of results between symbiotic and aposymbiotic *M. persicae* aphids for ASAI suggest that insecticidal effects of ASAI are present in symbiotic and aposymbiotic insects, but the poorer overall survival of aposymbiotic insects makes estimation of relative effects difficult.

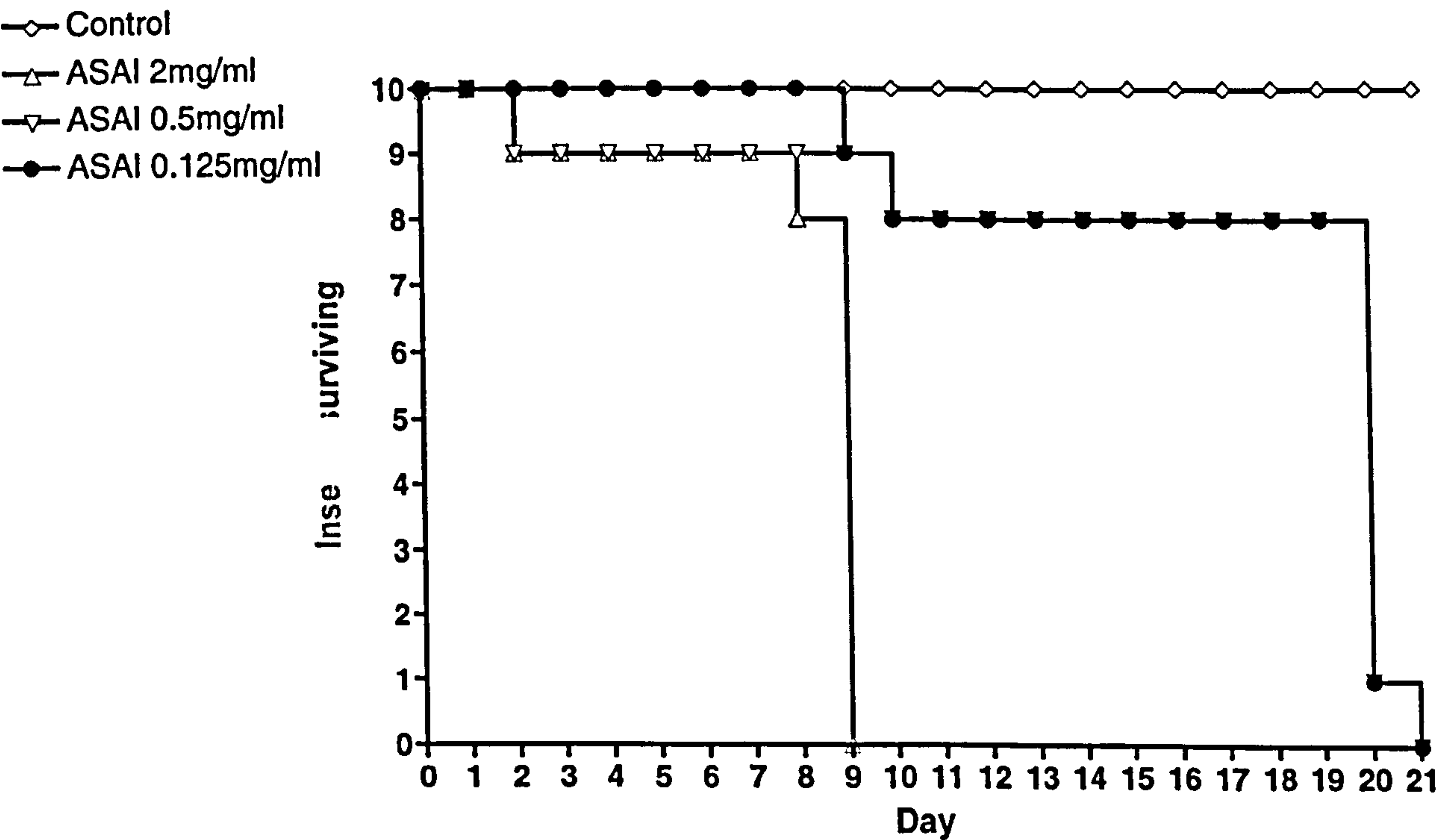


Figure 4.5 Artificial diet bioassay of ASAI against symbiotic *Myzus persicae*. Bioassay data supplied by Prof. Angela Douglas (University of York).

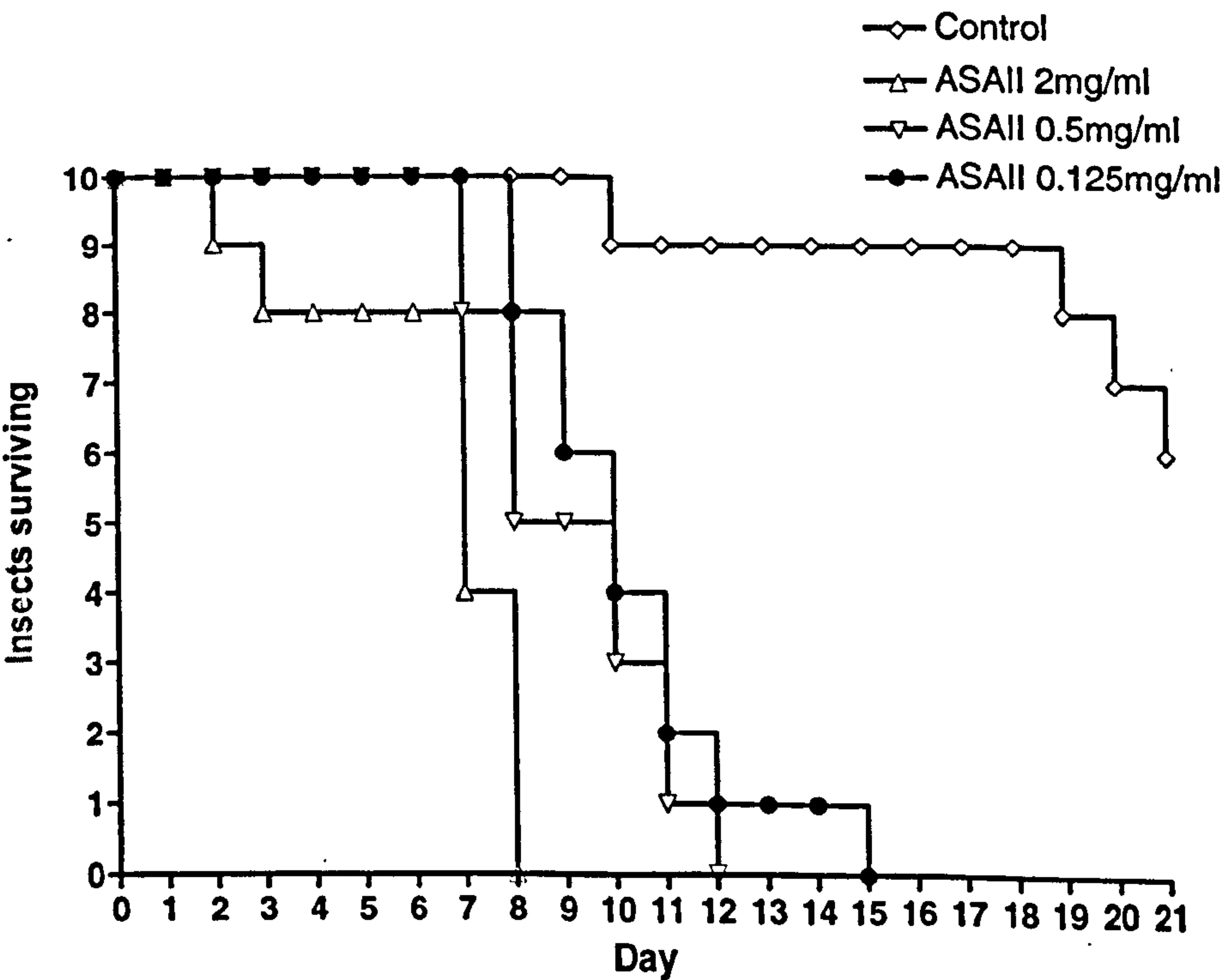


Figure 4.6 Artificial diet bioassay of ASAI against symbiotic *Myzus persicae*. Bioassay data supplied by Prof. Angela Douglas (University of York).

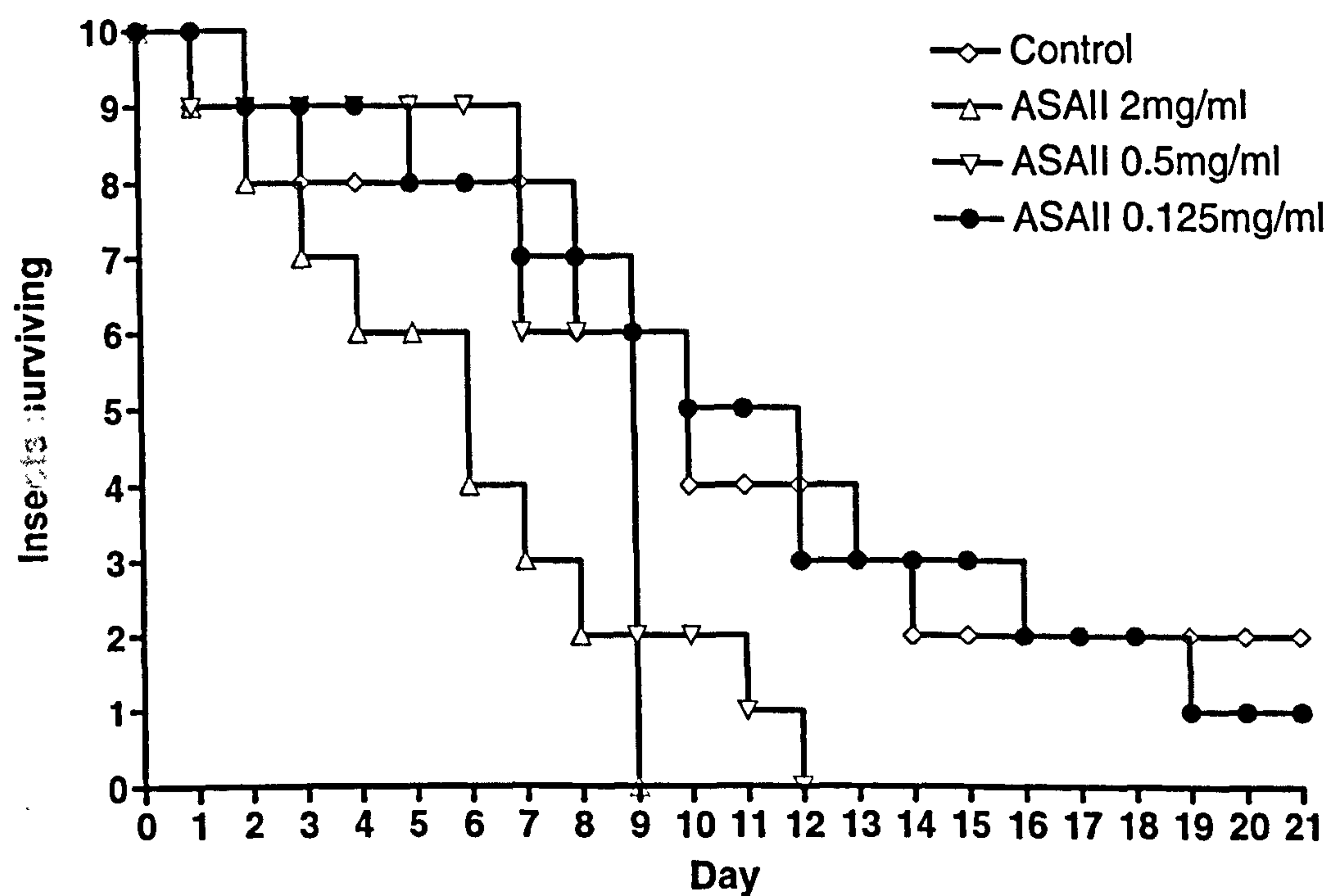


Figure 4.7 Artificial diet bioassay of ASaII against aposymbiotic *Myzus persicae*. Bioassay data supplied by Prof. Angela Douglas (University of York).

The failure to show any significant insecticidal effect on aposymbiotic aphids for the lowest concentration of ASAI tested (0.125 mg/ml) suggests insecticidal effects may be reduced in these insects compared to symbiotic aphids where an insecticidal effect was observed, but the poor survival of aposymbiotic aphids may prevent the rapid drop in survival after 20 days observed for symbiotic aphids being detected by the assay. Again, the overall conclusion is that there is no evidence to suggest that the insecticidal effects of garlic lectins are absent, or significantly reduced, in aposymbiotic aphids compared to symbiotic aphids.

4.2.3 Artificial Diet Bioassays Against Rice Brown Plant hopper (*Nilaparvata lugens*)

Artificial diet bioassays were carried out against third instar nymphs of *N. lugens*. The artificial diet system for these insects (MMD-1 diet) does not support survival and development as well as standard aphid diet supports *A. pisum* and *M. persicae*. Despite this, the insects could survive for a sufficient length of time on the MMD-1 diet to test recombinant ASAI and ASAI for toxic effects. ASAI marginally decreased survival of *N. lugens* at 1 mg/ml and 0.1 mg/ml but not in a dose-dependent manner. Differences in survival curves were not significant ($p > 0.05$, survival analysis), although the qualitative result was consistent. Decreased survival did not follow the no diet control, demonstrating that feeding was still occurring and recombinant lectins do not act as a strong feeding deterrent. Corrected mortalities (i.e. mortality relative to control at day when "no diet" insects were all dead) was 50% for both concentrations of ASAI tested. ASAI had a significant effect on survival ($p < 0.05$, survival analysis) when present at 1.0 mg/ml, but not when present at 0.1 mg/ml, when a marginal decrease in survival was observed. ASAI did not reduce survival to the level of the no diet control, so again feeding was not entirely eliminated. The corrected mortality values for ASAI were 32% at 0.1 mg/ml, and 82% at 1.0 mg/ml. Summaries of insect survival are shown in figures 4.8 and 4.9.

Mixed isoforms of garlic lectins purified from bulbs have previously been assayed against this insect and have been shown to exhibit insecticidal activity, and the results obtained with the recombinant ASAI are comparable to those assays. The results from artificial diet bioassays show that the recombinant lectins produced in *P. pastoris* retain the expected insecticidal properties towards hemipteran insects.

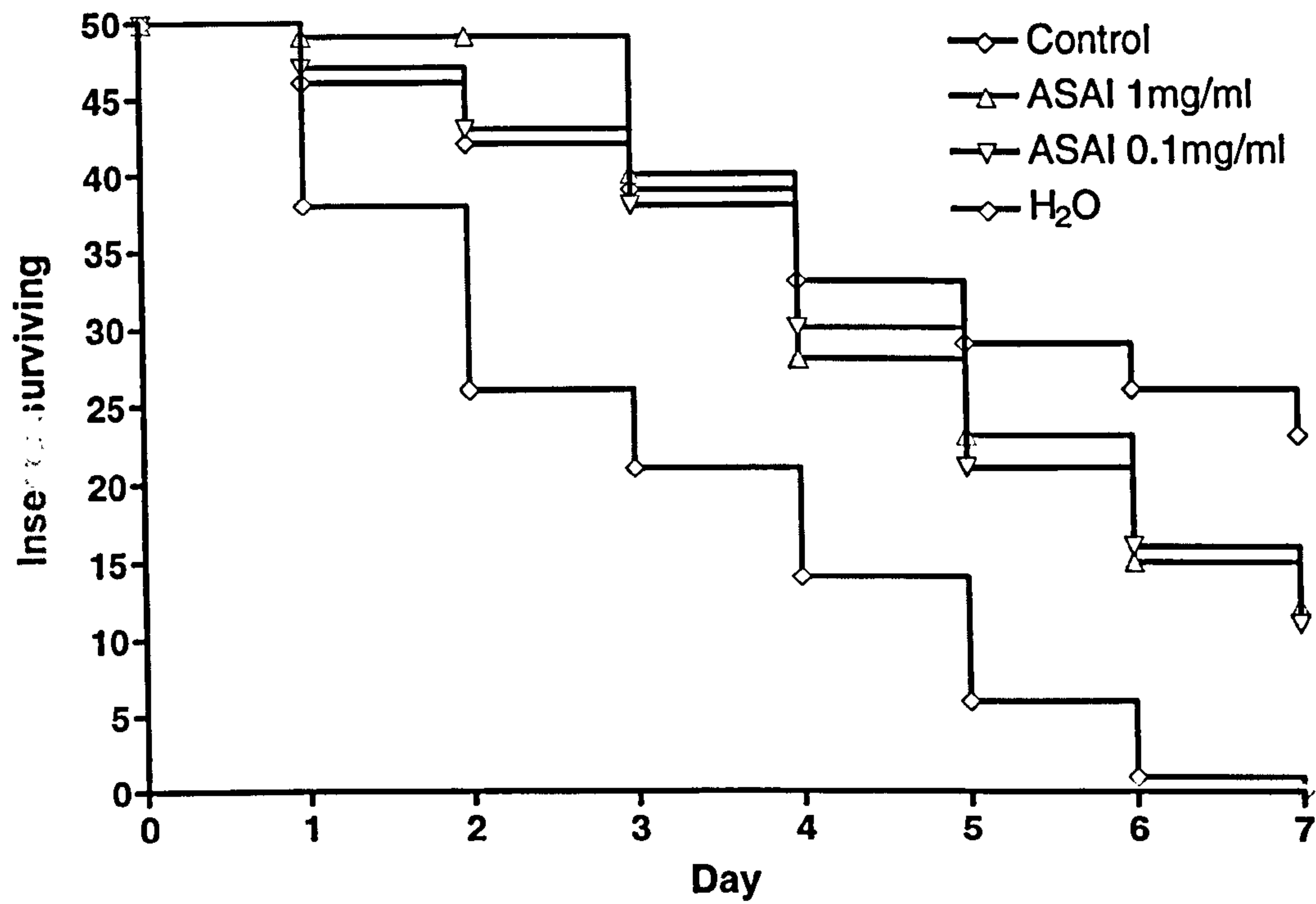


Figure 4.8 Artificial diet bioassay of ASAI against *Nilaparvata lugens*.

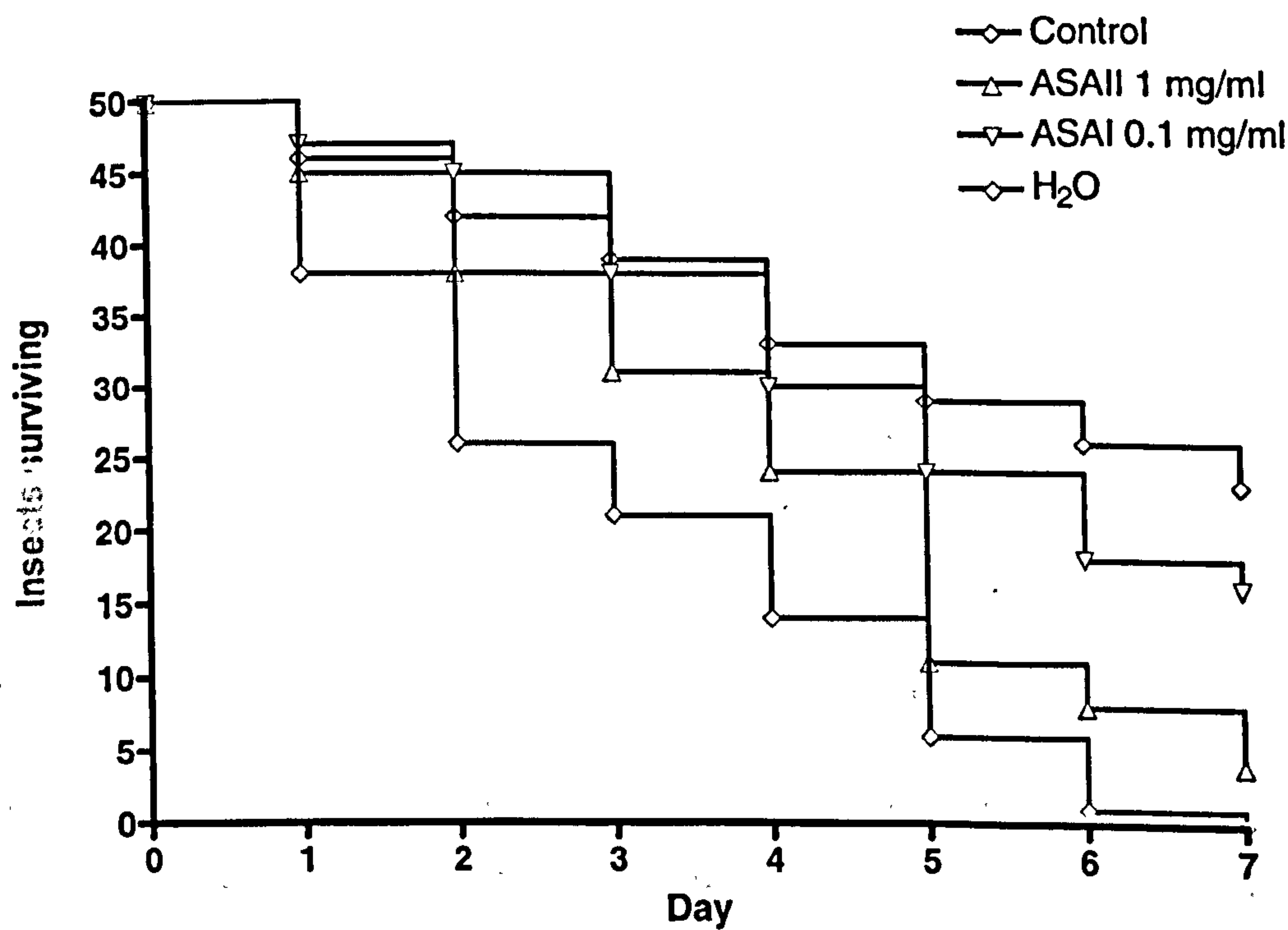


Figure 4.9 Artificial diet bioassay of ASAI against *Nilaparvata lugens*.



4.3 Characterisation of Lectin Binding to Insect Gut

In order to understand the toxicity of garlic lectins to hemipteran insects, it is important to characterise the interactions between the proteins and ligands on the surface of the insect gut that are potentially responsible for toxicity. The work of Banerjee *et al* (2004) showed that in mustard aphid (*Lipaphis erysimi*) garlic leaf lectin interacted with a *Buchnera*-produced protein, symbionin, and postulated that this interaction was important in the mode of toxicity. For this to be true symbionin must be released from the endosymbiont and localized to the aphid gut membrane. No evidence has been presented in the literature for this phenomenon. In addition, the interaction would not be expected to be mediated through the lectin's carbohydrate-binding activity, since symbionin is produced by the bacterial symbiont, which lacks a protein glycosylation system. To circumvent this problem, Banerjee *et al.* proposed that the binding of garlic lectin to symbionin was based on a protein-protein interaction, mediated by complementary hydrophobic surfaces on the two molecules.

In this study "receptors" for ASaII were investigated using a pull-down assay, in which immobilized ASaII was incubated with extracts from insect tissues, and MALDI-TOF mass spectroscopy. Initially, the assay was used on whole insect preparations, and once a protocol was established on aphid gut samples. Proteins present in the bound fraction of the pull down assay were analysed on SDS-PAGE and excised from the gel and subjected to tryptic digestion and MALDI-TOF analysis. The large number of *Acyrtosiphon pisum* ESTs present in bioinformatic databases were searched using the MALDI-TOF data obtained.

4.3.1 Isolation of Protein Fractions from *A. pisum*

Fractions corresponding to soluble and detergent-solubilized proteins from *Acyrtosiphon pisum* were prepared by differential extraction. Whole insects, or guts dissected in ice cold 0.9% NaCl were homogenized in 50mM HEPES buffer (pH 7.0). After centrifugation the supernatant contained the soluble protein fraction. The pellet was resuspended in 50mM HEPES containing 0.1% Triton X-100 and after further centrifugation the supernatant contained detergent-solubilised proteins, referred to subsequently as "insoluble". Triton X-100 treatment is sufficient to solubilise most

membrane proteins, but not many structural proteins. The two fractions had different band patterns when analysed by SDS-PAGE, although some cross-contamination in terms of bands present in both fractions was apparent.

4.3.2 Development of a Pull-Down Assay for Studying the Interaction of Recombinant Lectins with Aphid Gut Proteins

To study the interaction between ASaII and *A. pisum* proteins a pull-down assay using ASaII cross-linked to magnetic beads was developed. Cross-linking of 4 mg recombinant ASaII and 0.5 ml of amine-derivatized magnetic bead suspension was carried out using the water-soluble cross-linker BS³ in PBS (pH 7.0). According to spectrophotometric analysis at 280nm, 3.6 mg (90%) of recombinant ASaII was cross-linked successfully to magnetic beads (not shown), giving a concentration of 7.2 µg lectin per µl of bead suspension.

4.3.3 Pull-Down Assay of Whole *A. pisum* Insects

Soluble and insoluble protein fractions were prepared from approximately 90 mg of whole *A. pisum* insects as described earlier. Two pull-down assays were carried out, one on the soluble protein sample, the other on the insoluble protein sample, retaining samples of applied protein, unbound protein, wash fraction, and protein eluted from the beads by treatment with SDS sample buffer. In a control experiment with beads to which no protein had been attached, only minimal background binding was observed (result not shown). When proteins extracted from whole insects were tested, a large number of proteins in the soluble fraction bound to the immobilised ASaII, as shown by a complex and diverse band pattern for SDS-PAGE of the eluted fraction after silver staining (fig. 4.10). Specific binding had taken place, shown by the failure of some of the major components to bind, and their presence in the unbound proteins. The greater intensity of staining of the protein bands in the eluted fraction suggests that the proteins present had been concentrated by their interaction with the magnetic beads, implying that the pull-down was viable for a protein-protein interaction 'fishing' experiment. However, the amounts of protein bound by the magnetic beads could not be visualised by standard Coomassie blue staining after elution from the

magnetic beads, because the protein levels in the sample were too low. In the insoluble protein pull-down assay, more proteins appeared to be present in greater abundance than in the soluble pull-down. The majority of proteins in the insoluble fraction failed to bind, including most of the prominent bands, and were found in the non-bound fraction with a low amount of protein present in the eluted fraction. A band of approximately 50 kDa was the single dominant band in the eluted fraction. Overall, there was no carry-over of abundant proteins in the applied fraction into the eluted fraction, suggesting that the assay conditions were of a suitable stringency (fig. 4.10). The complexity of SDS-PAGE band patterns of the proteins bound by ASaII in the eluted soluble and insoluble protein fractions from whole insect made further characterisation of individual bands difficult or impossible, and required that the assay be carried out on proteins specifically derived from aphid guts.

4.3.4 Pull-down assay of proteins bound by ASaII in the aphid gut

With a suitable assay procedure set up it was possible to carry out a pull-down assay of aphid gut tissue only. In a preliminary experiment, approximately 50 aphid guts were dissected, and total gut proteins were extracted in HEPES buffer containing Triton X-100. The supernatant from this extraction was used for the pull-down assay. A sample corresponding to whole aphid protein (extracted with Triton X-100 containing buffer) was used as a comparison. Results are shown in figure 4.11.

The eluted fraction from gut proteins still shows a complex pattern on SDS-PAGE after silver staining, but major bands are evident, and bands are well-separated enough to make excision a viable possibility for the major components. The experimental strategy was considered to be viable for preparation of gels for subsequent MALDI-TOF analysis of protein components in the eluted fractions.

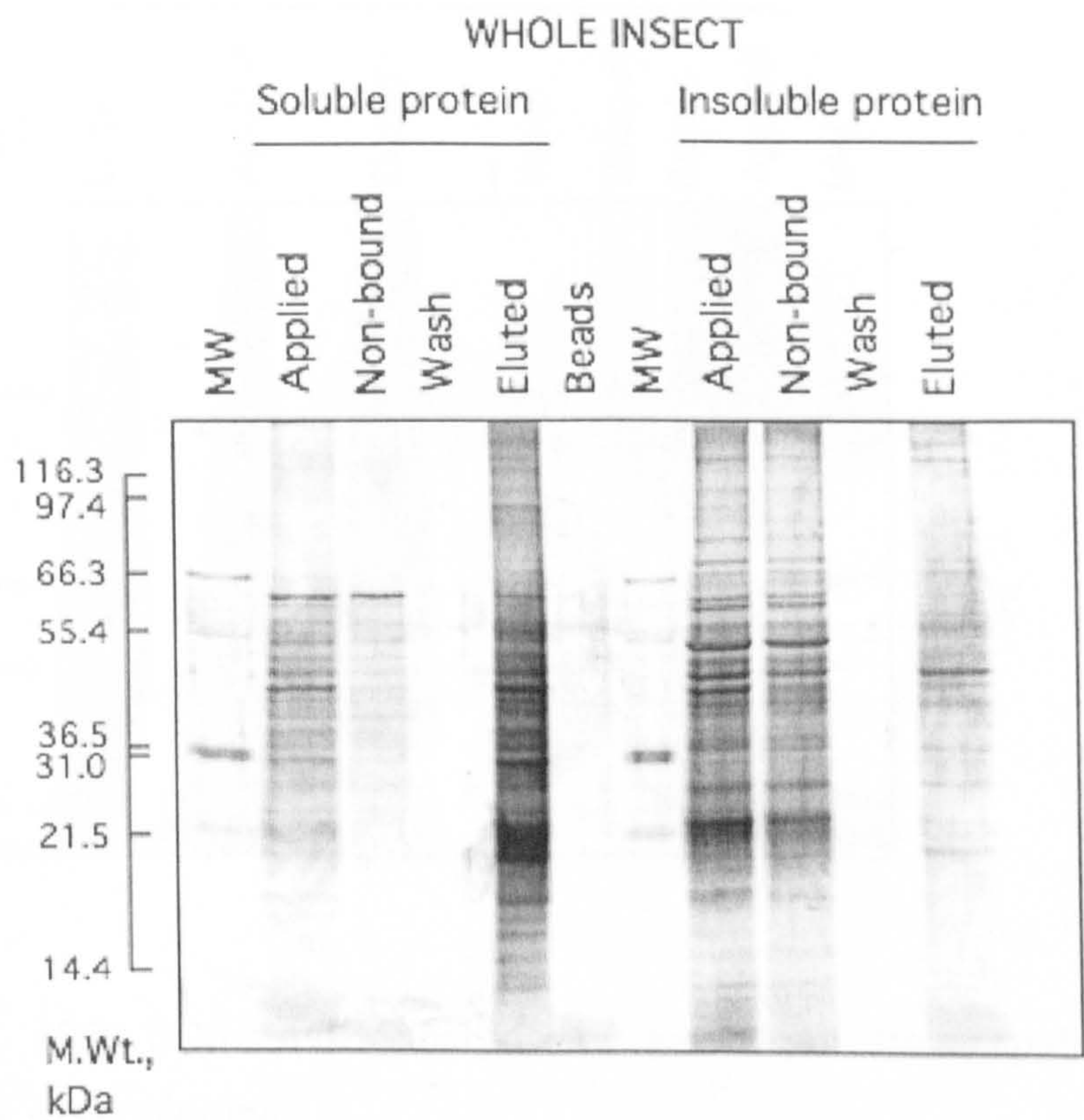


Figure 4.10 12.5% SDS-PAGE analysis of whole insect pull-down assay. Visualized by silver staining.

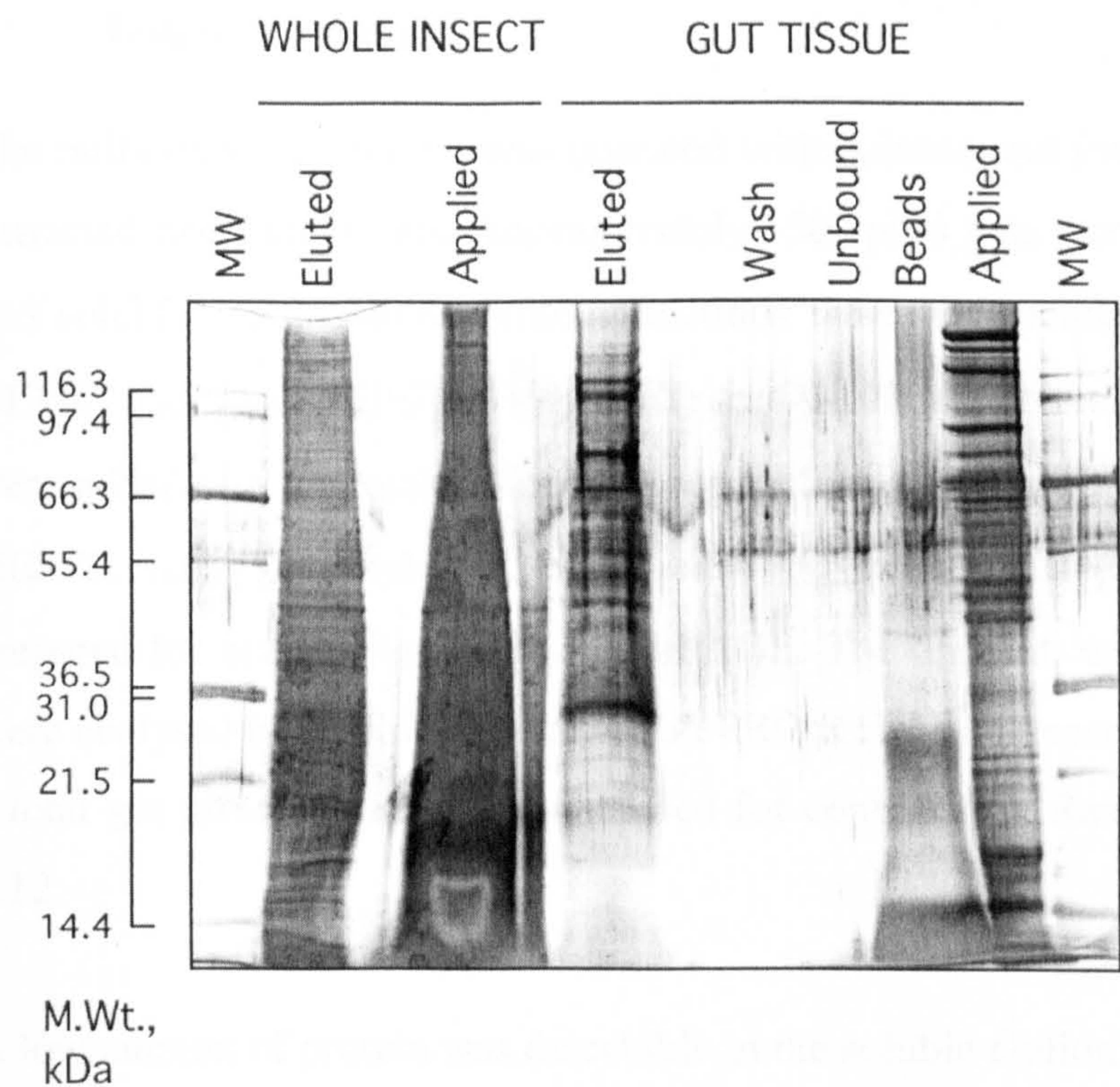


Figure 4.11 10% SDS-PAGE analysis of insect gut tissue pull-down assay. Visualized by silver staining.

4.3.5 MALDI-TOF mass spectrometric analysis of aphid gut proteins bound by ASaII

The pull-down experiment was repeated with soluble and insoluble protein fractions extracted from aphid guts. Approximately 150 aphid guts were dissected as described and soluble and insoluble protein fractions were resuspended in 200 µl of 50 mM HEPES and 50 mM HEPES + 0.1 % Triton X-100 respectively. 50 µl of ASaII beads were added and the total volume of the reaction was made up to 0.5 ml with 50 mM HEPES. After overnight incubation at 4 °C the beads were washed and samples prepared for analysis on 12.5% SDS-PAGE. The applied, wash, and eluted fractions were analysed by SDS-PAGE, using SYPRO Ruby fluorescent stain to identify bands; a total gut protein sample was included for comparison. Results are shown in figure 4.12.

A low amount of protein was detectable in the soluble elution fraction, the major band being at 66 kDa with lower molecular weight bands at approximately 55 kDa and 32 kDa. Amounts of proteins were too low for further analysis.

The elution fraction of the insoluble protein fraction pull-down contained approximately ten distinct protein bands, the most abundant being at approximately 110 kDa and 66 kDa. The band at 66 kDa is broad, and may represent a mixture of proteins; this band corresponds to the most abundant band in the total protein sample. The band at approximately 110kDa corresponds to an abundant band in the total protein sample also. These two bands were excised from the gel and subjected to MALDI-TOF analysis as they appeared to be retained most strongly in the eluted bead fraction.

Protein from the 110 and 66 kDa bands was extracted and digested with trypsin prior to analysis. MALDI-TOF analysis of the 110 kDa band produced a peptide mass fingerprint (PMF) (fig. 4.13) which was used in a Mascot search against the NCBI nr database. This gave a significant match to membrane alanyl aminopeptidase N (APN) from *A. pisum* (gi90656783) with a Mowse score of 143 where the database used gave a significance level of $p < 0.05$ for Mowse scores greater than 79. Membrane alanyl-APN had a molecular mass of 111198 Da in the sequence database and the PMF obtained had 33% sequence coverage (fig. 4.14). Only a single membrane-bound

aminopeptidase N occurs in *A. pisum* and accounts for around 15% of total gut protein. The enzyme is a glycoprotein that is rich in mannose residues and is known to bind the mannose and glucose specific lectin Concanavalin A (Con A) (Cristofolletti, P. T., *et al.*, 2006, Sauvion, N., *et al.*, 2004). The 66 kDa PMF (fig. 4.15) did not give any Mowse scores above 79; it did however give a match to sucrase from *A. pisum* (gi81159208) with a Mowse score of 66 and PMF sequence coverage of 30% (fig. 4.16). By increasing the peptide mass tolerance of the Mascot search from ± 50 ppm to ± 70 ppm a MOWSE score threshold of 73 was set and under this stringency sucrase scored 73, making it a significant match. The molecular mass of sucrase is 68845 Da and it was the only *A. pisum* sequence to be identified from the database. Recombinant sucrase has been shown to tightly associate to membranes in *Xenopus* oocytes (Price, D. R. G., *et al.*, 2007). Sucrase contains three potential N-glycosylation sites (N-X-S/T) but did not bind Con A agarose matrix (Dr. D. Price, personal communication). The mannose oligosaccharide specificity of Con A and ASaII may vary sufficiently for sucrase to bind selectively to ASaII but not Con A. Other sequences were identified in the Mascot search including predicted proteins similar to ventricular myosin heavy chain isoforms 1-7 from *Danio rerio* (Isoform 6 (gi68355960) had a Mowse score of 75) but these were from non-arthropod species, and had molecular masses inconsistent with the band on SDS-PAGE.

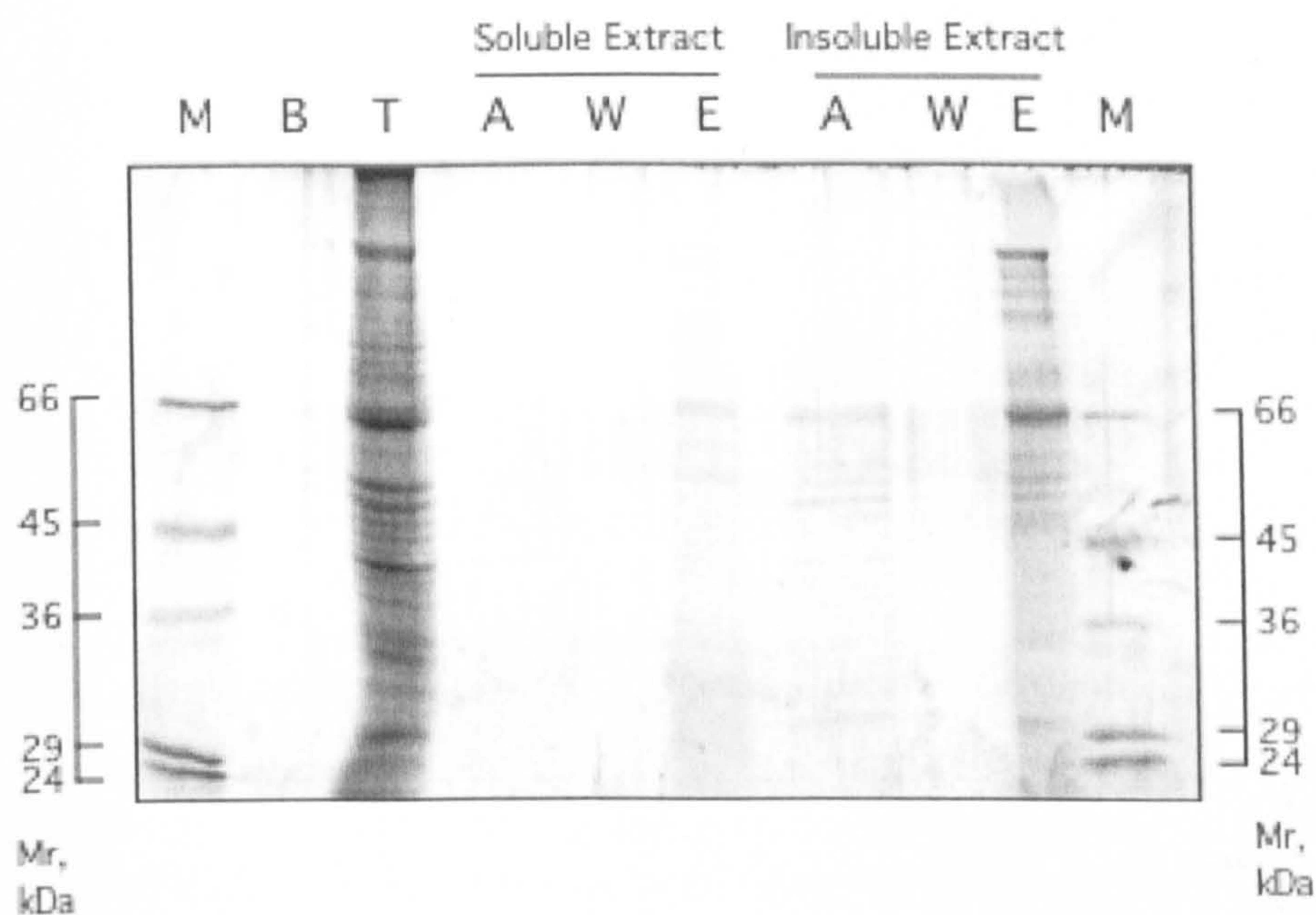


Figure 4.12 10% SDS-PAGE of *Acyrthosiphon pisum* gut protein pull-down assay. M = SDS7 marker. B = Bead only. T = Total gut protein. A = Applied sample. W = Wash fraction. E = Eluted, bound proteins. Visualized with SYPRO Ruby stain on a Typhoon imager (Amersham).

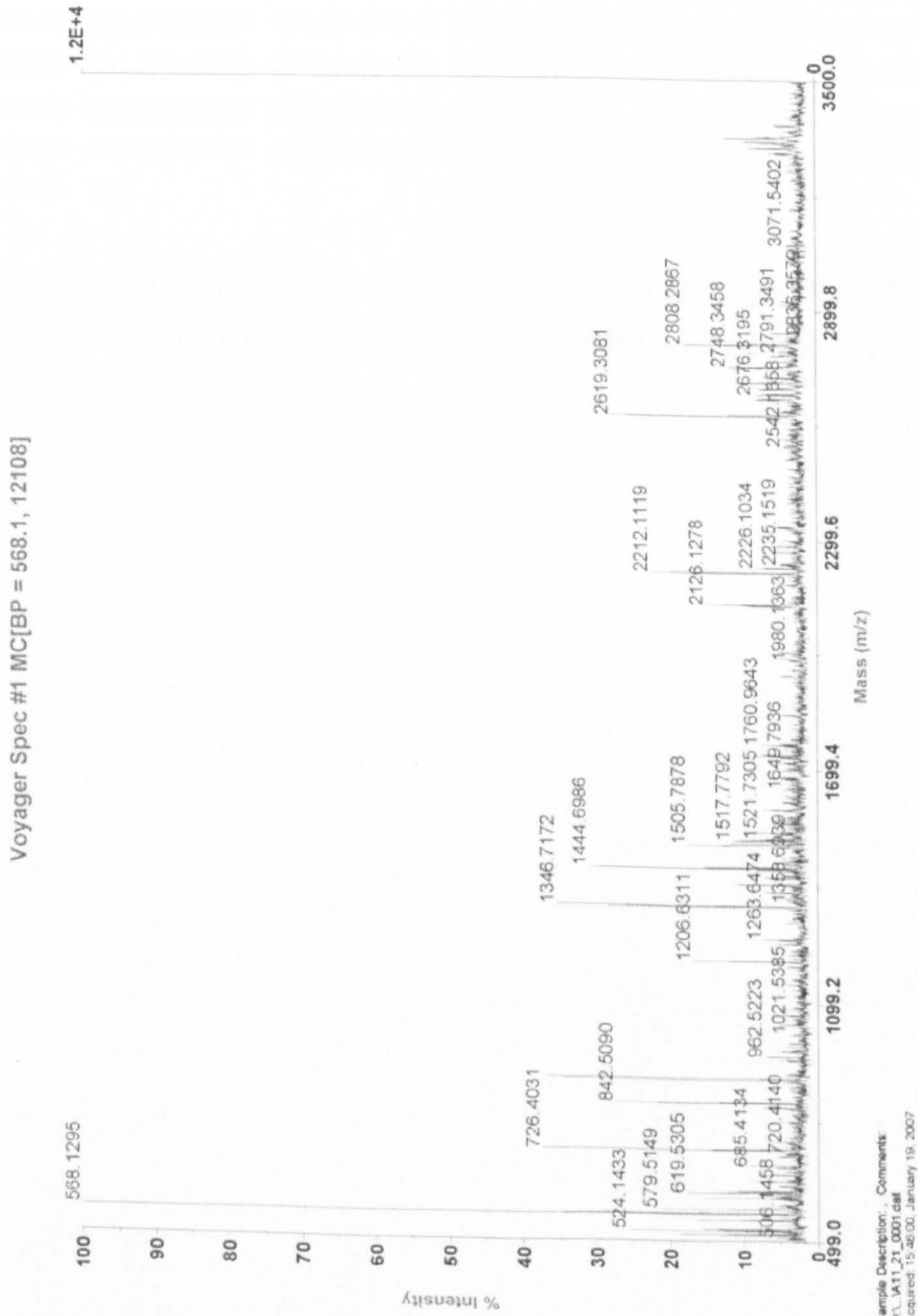


Figure 4.13 MALDI-TOF PMF for the 110 kDa band.

```

1  MFKLVAVLSA ALLCALVRGQ EDADDKKMDV FRLPENTSPE SYDLWFAPNM NDWTFEGCAK
61  ILVNINTPDT IAVTLNLNNL TVTNVSATDV SNNRDMVVAG LEYQTKNEQF VIRFOKAVPK
121 DRQLLVTIKY KGYIRDDNTG LYRSSYIEDG VTKWLAVTQF EPTSARLAFP CYDEPMYKAK
181 FNITVVKQNG QTVLSNMPIL KIEEGSKNTT VYFKETPPMS TYLAAIYVGE FVPKKNDSKI
241 TIITYKGKQG QTEYVATEAP KHLKVLEKYT GINYMLPKMD LLAIPDFRAG AMENWGMNTY
301 KESLLLLSND SKTKIKIQSS EIVQHEFTHQ WFGNLVTCKW WDYLWLNEGF AAYFQYFATG
361 MVRTSWPMEE LFLIDEFQNS MAYDQTPRHP ITTSVNTPQE TVDIFDAITY RKAASVLRML
421 RYTVTENLFO LSLOYLLKQF SYKAADPTNL FSAFDNVLYD KNYEIGNGLT VNEFMSNWTL
481 QSGYPVLNVT RNATTNTFSV TQRQFYVNKT DTTINETAAW HICLTYITDV SKNFTYNSSS
541 SVWTNKESPL TVFEAPEDIN WYLENPKAIG FYRVNYDVNN WMALIKQLNE TPTDIHVLSR
601 AQLIDDAFNL ARSGOLDYSV ALHLSKYLKY ENNTTPWYSA MNVFSYLLER MPRSDEGYED
661 FKSHVSSLAG TIYKKLEDLV SSGNTEFKVL SAWETFSSWA CRLENKYCLT KAQEYFNKWQ
721 AGEKIPADIK DAAFCIGVKN SNGSSVFNNM LALYNSTNSI SEKSSAESAL GCSTNRTOLS
781 EYINHILDGP NGLIKYKDMR SVVSAISVTP LGLDVLYEFM STNINITMDQ LPNGEDIITF
841 IYSTLASKVT NDDEIEKINN EKNNSTVRAS VRPMFENSYK VVEYNLEWFD SYSVGINQWT
901 SAYGEKEPQD IEITTLPPTT IAAGTTTGST SMNVTVKIST TTPKPSSATL NAYSSLVILQ
961 LTVFTVLIKT YFL

```

Figure 4.14 Amino acid sequence data of membrane alanyl aminopeptidase N (ABD96614.1) from *A. pisum*. Bold and underline typeface signifies peptides identified in mass spectrum (fig. 4.13).

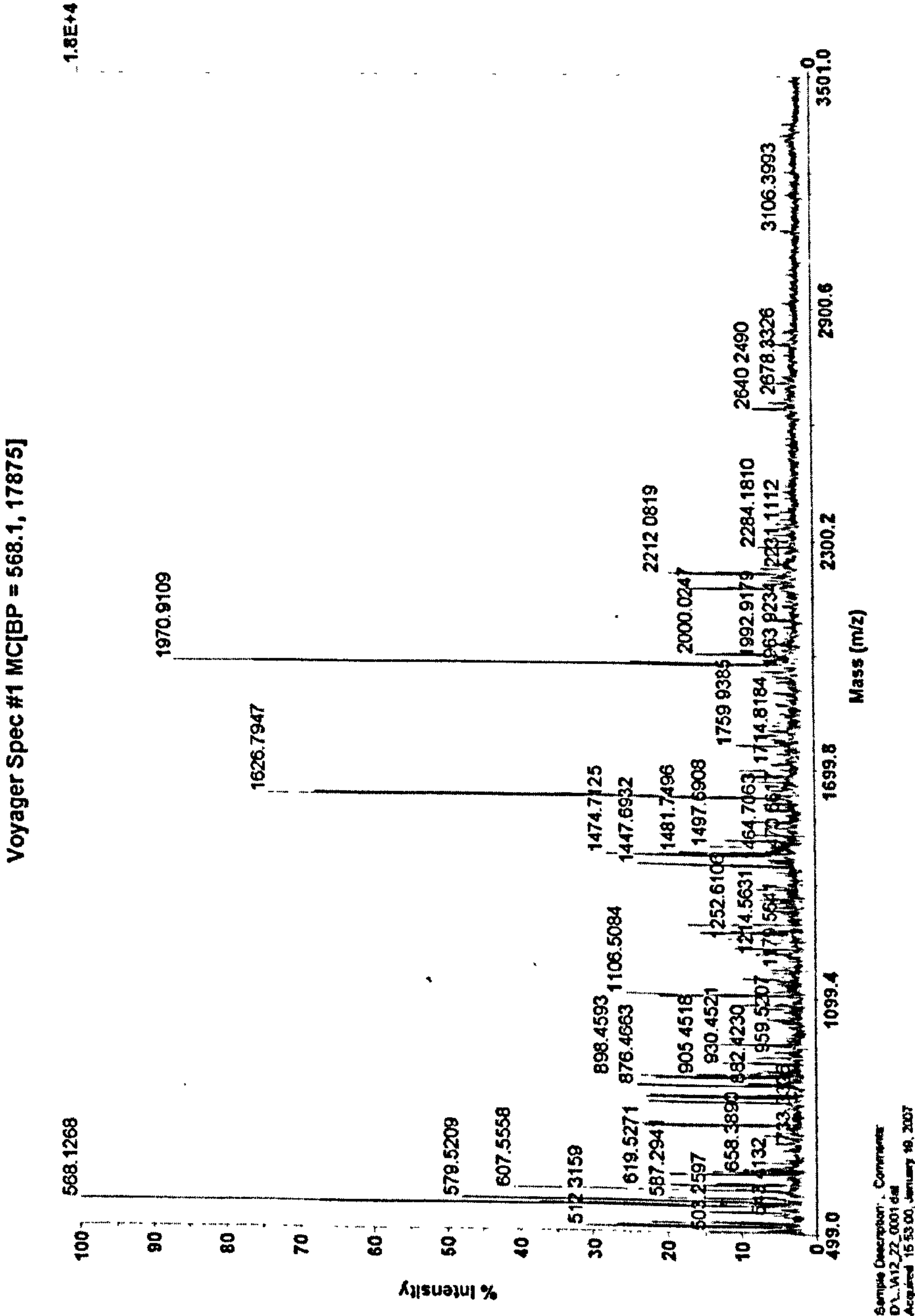


Figure 4.15 MALDI-TOF spectrum for 66 kDa band.

1 MFKVITVCLW LFAFNSLNVS SEYVYEG LKS DSVEPDWWQT EIIYQIYVRS **FKDS****SDGDGIG**
61 **DLNGITEKVP** **YFKTIDVGAV** WLSPIFLSPQ NDFGYDISDY KEIDPIYGSM ADFERM**RDEF**
121 **HKHGIKVLLD** **FVPNHTSDEH** **EWFOKSIKKI** EPFSDYYVWK DPIRDVHGNN TPPSNWLGVF
181 NSGSAWEWNE ERQQYYLHQF **QVKOPDLNYR** **NPSVREEIKN** **TLLYWLGRGV** DGFR**FDAVNY**
241 **LFEREDLADE** **PKSNKIGYLD** TDYDSLTHTS TLDQPETYSI VRQWRQMLDS YRTREKKTKF
301 MMVECYSFPD KTLLEYGSNS EPGAHF PFNF LFIGTFDQOS DAAK**VHNMIK** SWIR**GMPTGM**
361 **WPNWVLGNHD** **NARVASRSNP** MLVDGLHMIQ HLLPRTSVTY YGDELGLIDT TVR**WDQTVDP**
421 **AGLNVGPYRF** LK**FSRDPVRT** PFPWDSSYNA GFSNSSSLWL PLNADYWKKN MVEESRFKSN
481 LRSYRQLARL RRS LTFVKGD LHLYT LSK**WV** **FGFSRSFYDH** PTYFIVVNFG SEIETVNLME
541 ARGTLPLTMK VK**VSSINSGF** **VTGNLVRTDS** VLLRPKAALV LTTSRMNEDA

Figure 4.16 Predicted amino acid sequence of sucrase from *A. pisum*. Bold and underline typeface signifies peptides identified in mass spectrum (fig. 4.15).

4.4 Analysis of Lectin Transport into the Aphid Haemolymph

To determine whether or not recombinant garlic lectin was delivered to the haemolymph, day 0 *A. pisum* were fed recombinant lectin and Western blots were carried out on extracted haemolymph after 72 hours. The sensitivity of anti-ASA antibodies was not sufficient to detect the levels of lectin expected in the haemolymph (<10 ng), it was therefore necessary to use an ASaII-Avidin fusion protein with avidin only controls in order to detect intact ASaII-Avidin in the haemolymph using anti-avidin antibodies. Details of ASaII-Avidin are given in chapter 5. 50 insects were fed on control diet and 50 were fed on ASaII-Avidin at a concentration of 0.25 mg/ml and all insects were bled after 96 hours feeding. Proteins present in the haemolymph were separated by SDS-PAGE and probed by Western blotting using anti-avidin antibodies. Western blot analysis showed the presence of a 20 kDa protein in the haemolymph of ASaII-Avidin fed insects which corresponds to the molecular weight of the avidin subunit alone. No intact ASaII-Avidin (~36 kDa) was detected in the aphid haemolymph (fig. 4.17). A similar Western blot carried out on insects fed avidin alone showed that avidin is transported across the aphid gut without ASaII, suggesting that ASaII is not necessary for avidin to be transported (fig. 4.18). It is not possible to assert that ASaII is transported across the aphid haemolymph from this data.

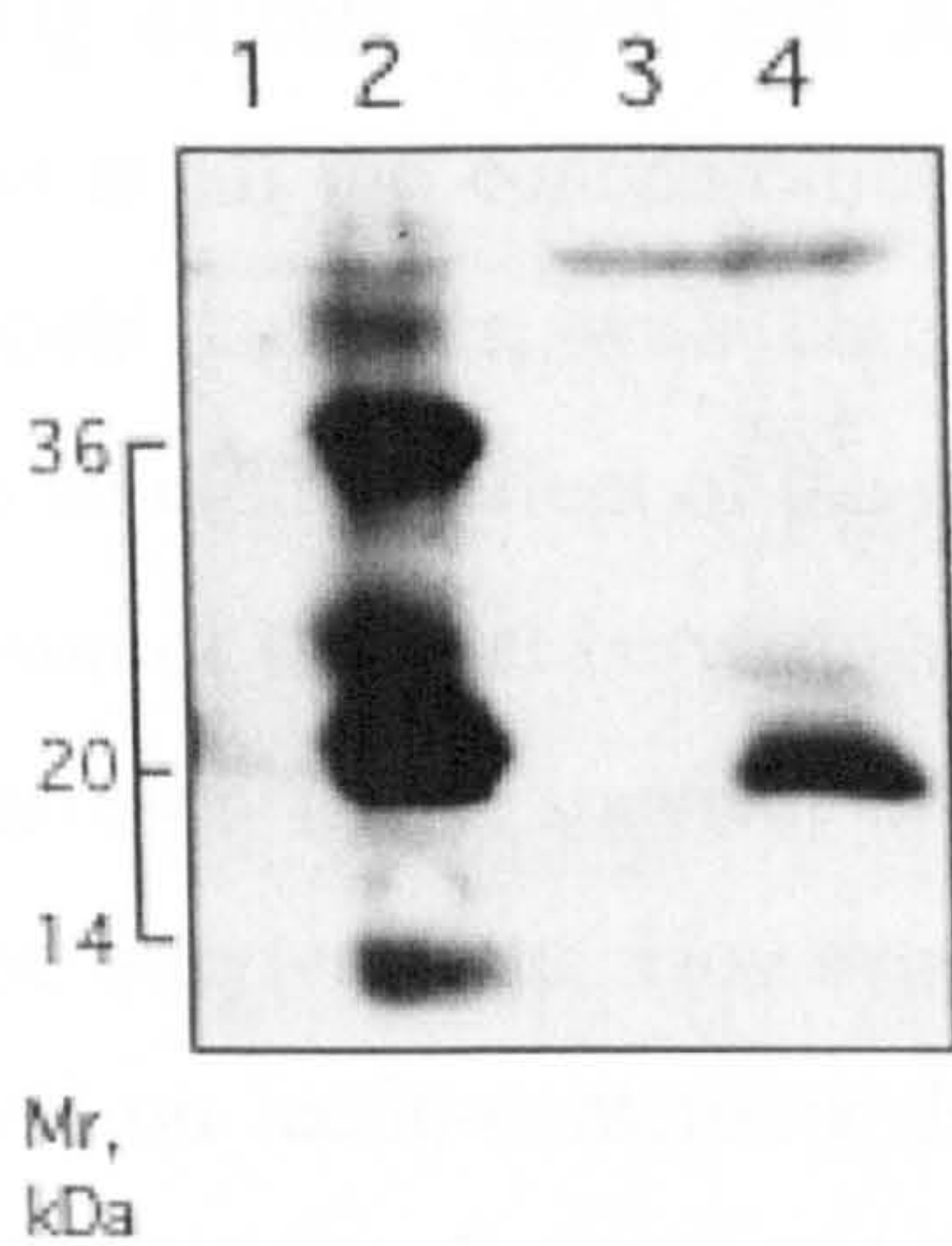


Figure 4.17 Western blot showing (1) control diet, (2) 0.25 mg/ml ASaII-Avidin in diet, (3) haemolymph from control fed insects after 96 hours and (4) haemolymph from ASaII-Avidin diet fed insects after 96 hours. Detected using anti-avidin antibodies (1:10000).

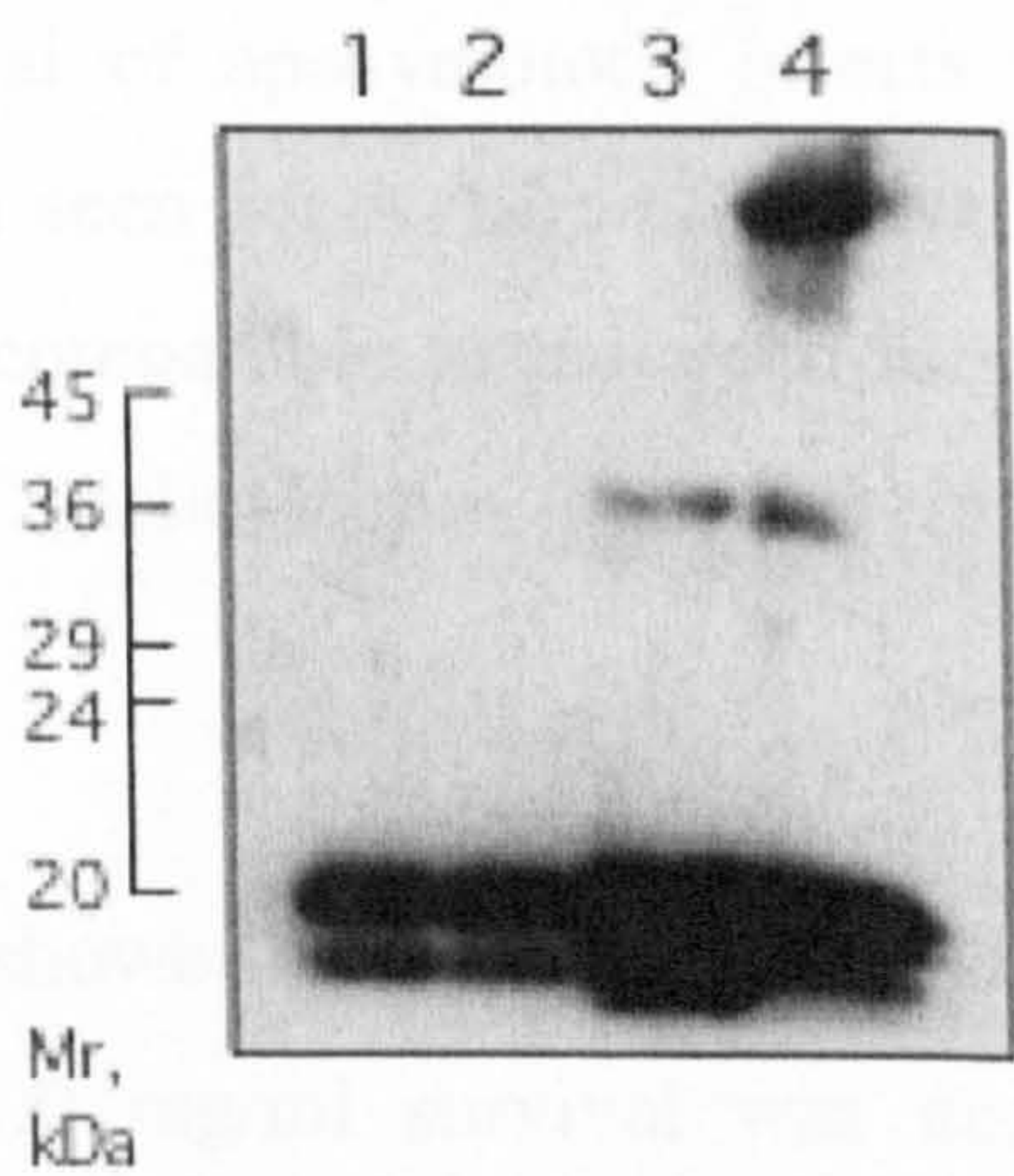


Figure 4.18 Western blot showing (1) avidin standard (25 ng), (2) avidin standard (50 ng), (3) 0.25 mg/ml avidin containing diet and (4) haemolymph from avidin diet fed insects after 96 hours. Detected using anti-avidin antibodies (1:10000).

4.5 Discussion

Artificial diet feeding assays have demonstrated that both recombinant ASAI and ASAII were toxic to the hemipteran insects *Acyrtosiphon pisum*, *Myzus persicae* and *Nilaparvata lugens* to differing extents. ASAI and ASAII showed a significantly deleterious effect on *A. pisum* at all the concentrations tested in a dose-dependent manner. Toxicity was not acute however, with chronic effects seen past day 4, suggesting toxicity was due to an indirect effect of the recombinant lectins rather than a direct toxic effect. Comparison of survival between symbiotic and aposymbiotic *A. pisum* was complicated by variation in the survival of control insects, aposymbiotic insects displaying a decreased survival rate. However, contrary to the findings of Banerjee *et. al* (2004) survival on lectin-containing diets was comparable between symbiotic and aposymbiotic insects. It was suggested in this research that aposymbiotic insects would demonstrate increased survival when fed only lectin-containing diets on the basis of the binding interaction between the bacterially produced symbionin protein and garlic leaf lectin in the mustard aphid (*L. erysimi*).

A similar pattern of toxicity was seen in the artificial diet bioassays against *M. persicae*. Both ASAI and ASAII had significant effects on the survival of symbiotic insects, with the highest concentration 2 mg/ml demonstrating the greatest level of toxicity. Once again survival of aposymbiotic insects on control diet was highly variable, even more so than seen for *A. pisum*, survival of aposymbiotic insects on lectin-containing diets was comparable to that seen in symbiotic insects, suggesting further that the toxicity of garlic lectins does not involve an interaction with a symbiont-produced protein.

The results from *N. lugens* showed that ASAI did not significantly affect survival of the insect but ASAII at 1.0 mg/ml survival was decreased significantly. These findings were in line with previous assays of mixed garlic lectin isoforms. *N. lugens* does not possess *Buchnera*-like endosymbionts but some toxicity is still observed suggesting that the mode of action is independent of bacterial endosymbionts.

It was noticeable that ASAII demonstrated a greater toxicity than ASAI towards all hemipteran insect species tested. This may reflect the ability of ASAII to form stable

homodimers from single subunits spontaneously, resulting in a molecule with a greater mannose binding capacity than ASAI which cannot form dimers spontaneously as both subunits are encoded in its amino acid sequence. Variable proteolytic cleavage of the ASAI heterodimer has been described in chapter 3 and the absence of intact heterodimers may contribute to a decreased level of toxicity compared to ASAI.

To further understand the mode of action of lectin toxicity and lectin transport across the aphid gut a pull-down assay was successfully developed to isolate garlic lectin-binding proteins within the aphid gut. This assay identified two gut proteins with significant Mowse values after MALDI-TOF analysis: Alanyl-APN and sucrase.

A single membrane alanyl-APN occurs in *A. pisum* (Cristofolletti, P. T., *et al.* 2006) and it accounts for more than 15% of total gut protein and is rich in mannose residues which bind to recombinant ASAI. APN is present predominantly associated with the microvillar membrane of the insect midgut, approximately 98% of total APN is membrane-bound, the remaining 2% is present as a soluble enzyme. It is therefore perhaps unsurprising to observe an interaction between ASAI and APN in the insect midgut. APN has been established as the target of the insecticidal δ -endotoxins from *Bacillus thuringiensis* (*Bt* toxins). *Bt* toxins exert toxicity to insects by the formation of channels through the midgut. Interestingly no successful *Bt* toxin has been found against hemipteran insects.

Sucrase is a predominantly membrane bound enzyme integral to osmotic regulation in the aphid gut. The observation that sucrase was unable to bind to Con A suggests that it does not possess mannose oligosaccharide subunits despite it being a glycosylated protein. It may therefore be the case that a protein-protein interaction with ASAI is occurring.

It was not possible to determine whether or not recombinant ASAI was transported across the aphid gut into the haemolymph by Western blotting. Transport of other lectins across the gut of hemipteran insects, for example GNA has been detected in

the haemolymph of *N. lugens* (Down, R. E., *et al.*, 2006). Some toxicity towards *N. lugens* was observed in this paper and transport to the insect haemolymph was implicated in this toxicity. Transport of a spider venom peptide as part of a fusion protein across the insect gut was also achieved and this enabled the toxic effects of the venom peptide to greatly increase toxicity. In this study involving garlic lectin, it was not possible to show intact ASaII-Avidin in the aphid haemolymph, only the avidin subunit. This may suggest that ASaII can transport avidin across the aphid gut, but it has also been demonstrated that avidin alone can cross the gut. It may be the case that ASaII-Avidin is cleaved within the aphid gut and the avidin subunit is preferentially transported. Without knowledge of the fate of ASaII after ingestion it is difficult to suggest how the toxicity observed in insect bioassays comes about. Data from MALDI-TOF has revealed that ASaII interacts strongly with APN and to lesser extent sucrase. Binding to these proteins in the gut membrane could produce physiologically significant effects on insect survival, particularly if amino acid absorption or breakdown of carbohydrate molecules is significantly inhibited.

Chapter 5

Development of Insecticidal Fusion Proteins Based on Garlic Lectin

5.1 Introduction

The stability of snowdrop lectin (*Galanthus nivalis* agglutinin; GNA) to gut proteolysis in insects and its ability to cross the gut into the haemolymph when orally administered has been demonstrated (Fitches, E. C. *et al.*, 2001). Subsequently GNA has been employed as a “carrier” protein for potentially insecticidal peptides (Fitches, E. C. *et al.*, 2002) and proteins (Fitches, E. C. *et al.*, 2004) in recombinant fusion proteins. Since garlic lectins have a similar binding specificity to snowdrop lectins, and their similarity in sequence suggests they are also able to resist proteolysis, these proteins were expected to transport across the insect gut into the haemolymph like GNA. On this basis, fusion proteins using recombinant ASaII as an alternative “carrier” protein were created and tested against insect pests. Since ASaII is present in a food crop that is consumed raw, there would be fewer potential objections against its use in an insecticidal treatment. Two strategies were employed for the production of fusion proteins based on ASaII: in the first, the methods used to produce GNA-based recombinant fusion proteins, containing the lectin polypeptide fused to an insecticidal protein toxin, were copied, with suitable adaptations; in the second, a recombinant ASaII-Avidin fusion protein, which would allow linkage of biotinylated active peptides to the carrier, was produced.

5.1.1 Insecticidal fusion proteins containing a toxin and a lectin carrier

In order to produce insecticidal fusion proteins based on GNA, it was necessary to establish sources of potentially insecticidal proteins or peptides. Venoms of predatory arthropods are a rich source of insecticidal molecules, many of which are insect-specific toxins. Previously, a gene encoding a venom component from the arachnid species *Segestria florentina* (SfII I) was obtained and incorporated into an expression construct used to produce a SfII-GNA fusion protein (Fitches, E. C. *et al.*, 2004). Subsequently, the

venom of the red scorpion *Mesobuthus tamulus*, family *Buthidae*, was used as a source of toxic components. When initially isolated and characterized, the red scorpion toxin (RST) was described as putatively lepidopteran-specific (Wudayagiri, R. *et al.*, 2001) but artificial diet feeding trials of a RST-GNA fusion protein demonstrated toxicity to lepidopteran and hemipteran insects (Pham Trung, N. *et al.*, 2006). The expression construct for RST-GNA in this case contained a synthetic RST coding sequence. The toxin itself demonstrates a high degree of sequence similarity to other insect toxins such as neurotoxin P2 and chlorotoxin. The RST scorpion toxin was selected for incorporation into fusion proteins based on ASaII as a carrier.

5.1.2 Insecticidal fusion proteins based on a carrier to which biotinylated peptides can be attached

Another means of producing insecticidal fusion-proteins, particularly those which utilize potentially insecticidal peptides as toxic agents, is via a system which allows chemically-synthesised peptides to be attached to a carrier. Such a system would provide a convenient way to test a number of potentially insecticidal peptides without having to produce numerous expression constructs. In addition, it would allow peptides containing modified amino acids to be used in fusion proteins. Post-translational modification of recombinant proteins by expression systems is determined by the host, and yeast expression systems generally lack the capacity to produce the kinds of modification that are common in insect peptide hormones and toxins, such as C-terminal amidation.

A system for producing peptide-carrier conjugates developed in Durham was chosen for this project. It is based on a recombinant fusion protein containing ASaII fused to the N-terminus of chicken egg white avidin (ASaII-Avidin). In this fusion, the ASaII acts as a “carrier” component, to allow transport across the insect gut, and the avidin acts as a binding component, to which biotinylated molecules can be bound. The strong binding between avidin and biotin (K_d approximately 10^{-15} M; (Kohanski, R. A., Lane, M. D., 1990)) would make the formation of complexes effectively irreversible. Biotinylated peptides can be chemically synthesised, and conjugated to a recombinant ASaII-Avidin

fusion protein. This system allows almost any peptide that has been characterised at the amino acid level to be exploited as a potential insecticidal toxin. Besides using peptides from venoms, endogenous insect peptide hormones can also be employed as potential toxins, since exogenously supplied hormones can disrupt insect physiology and development. The carrier system allows these peptides to be delivered orally, to reach their sites of action, whereas normally they would only be effective if injected into the insect haemolymph.

To evaluate the ASaII-Avidin fusion protein as a peptide delivery vehicle, an endogenous insect peptide hormone was selected as the potential toxin to be coupled to the carrier. Leucomyosuppressin (LMS) is an FLRFamide peptide of the FMRFamide-related peptide (FaRP) family. FaRPs occur widely across the animal kingdom and possess either the tetrapeptide FMRFamide or structurally similar FLRFamide C-terminal peptide sequence. In insects FLRFamides primarily show strong myotropic effects by inhibition of visceral muscle contraction, and consequently have become known as myosuppressins, although other effects including stimulation of skeletal muscle contraction have been observed (Kingan *et al.*, 1990). The consensus leucomyosuppressin sequence XDVXHXFLRFamide has been characterised in a number of insect species. These include the cockroaches *Leucophaea maderae* (Holman *et al.*, 1986) and *Diploptera punctata* (Donly *et al.*, 1996), the locusts *Locusta migratoria* (Robb *et al.*, 1989) and *Schistocerca gregaria* (Schoofs *et al.*, 1993), the dipteran species *Drosophila melanogaster* (Nichols, 1992) and *Neobellieria bullata* (Fonagy *et al.*, 1992) and a lepidopteran species, *Manduca sexta* (Kingan *et al.*, 1990). In *D. punctata* LMS mRNAs have only been detected in the anterior gut region in apparent endocrine cells, although immunohistochemical analysis shows FMRFamide-like immunoreactivity in posterior endocrine cells and nerve tracts along the entire gut.

No LMS-like peptides have been characterized in hemipteran insects but the fact that FMRFamide peptides are so widespread across the order *Insecta* suggests that it is likely that hemipteran insects, including aphids, will possess a similar peptide. The importance of regulation of gut contractions is important in aphids as it is in other insect species;

although aphids do not have to deal with protein-rich food requiring enzymatic digestion, they do have to exert tight control over ingestion of large volumes of liquid phloem sap. A disruption to this ability is potentially fatal to the aphid.

5.2 ASaII-RST: A Recombinant Fusion Protein of Garlic Lectin and Red Scorpion Toxin

The synthetic gene sequence encoding a mature toxin from *Mesobuthus tamulus* (red scorpion toxin, RST) described by Pham Trung *et al.* (2006) was used as the basis for a construct in which this toxin was fused to the garlic lectin ASaII as a carrier. The fusion protein described by Pham Trung *et al.* (2006) would be designated RST-GNA in the nomenclature used in this thesis (referred to as ButaIT-GNA in the paper; ButaIT was used as an abbreviation instead of RST to agree with earlier literature). However, on the basis of variants of the basic fusion protein between GNA and RST produced in the laboratory, it was decided to produce a fusion protein based on ASaII with the components in the opposite order, which was designated ASaII-RST. Two constructs were produced for the expression of ASaII-RST. Initially a construct containing full-length ASaII fused N-terminally to the RST coding sequence was produced (ASaII₁₂₇-RST). A second construct was designed in light of the information obtained from MALDI-TOF analysis of ASaII, which showed that C-terminal cleavage of ASaII took place in *Pichia pastoris* during synthesis. This revised construct (ASaII₁₁₅-RST) encoded the first 115 amino acids of ASaII (up to the predicted C-terminal cleavage site) fused N-terminally to the RST coding sequence. ASaII-RST fusion proteins were produced by Dr. Judith Philip.

5.2.1 The ASaII₁₂₇-RST construct

The construct containing a “full-length” ASaII coding sequence, ASaII₁₂₇-RST was assembled as follows. The RST fragment was PCR-amplified from an existing construct

containing GNA-RST, to produce RST with a 5' *Not* I site and a 3' *Xba* I site, using the following oligonucleotides (restriction sites underlined):

RSTfwd: 5' - TAAGCGGCCGCAAGGTGTGGTCCTTG - 3'

RSTrev: 5' - TAATCTAGATCATTGTATACCACAGA - 3'

The amplified RST fragment was separated by agarose gel electrophoresis, purified, and sequentially digested with *Xba* I and *Not* I. It was then ligated to the ASAIIPGAPzαA plasmid which had also been digested with *Xba* I and *Not* I. The ligation mixture was transformed into *E. coli*, and colonies containing the required recombinant vector were selected by plating on media containing zeocin. Plasmid minipreps were carried out on putative recombinants, and plasmid DNA was analysed by restriction digestion, and by DNA sequencing. The resulting clone contained the correct product for RST ligation into ASAIIPGAPzαA. The sequence for ASAII₁₂₇-RST is shown in figure 5.1. The correct construct was transformed into protease A-deficient *P. pastoris* strain SMD1168, and transformants were selected by plating on zeocin-containing media. Selected colonies were picked off and cultured in a small-scale shake flask experiment. Culture supernatants were screened for expression of the recombinant fusion protein by Western blotting, as described for recombinant garlic lectins.

Screening of supernatants from small-scale cultures showed that the fusion protein had undergone significant cleavage, and that although a band of the expected size for the ASAII₁₂₇-RST fusion protein (molecular weight approximately 20 kDa) was present on Western blots probed with anti-ASAII antibodies, much of the immunoreactive material was at the same molecular weight as the ASAII standard.

The ASAII₁₂₇-RST fusion protein was produced on a large scale by fermentation of a selected clone, as for recombinant garlic lectins. Protein was purified from culture supernatant by ion exchange chromatography purification on S-Sepharose, using a method similar to that carried out for the recombinant garlic lectins expressed in *P. pastoris*. Culture supernatant was diluted 10-fold into 50mM sodium acetate pH 4.0, and applied to the column, which was eluted with a linear gradient of [NaCl], from 0-1.0 M. ASAII₁₂₇-RST eluted as a broad peak, contaminated with yeast proteins; the recombinant

protein was further purified by gel filtration of a column of Sephacryl S-200. Analysis of purified ASaII₁₂₇-RST on SDS-PAGE showed that almost complete degradation (>90%) of the intact fusion had occurred (fig. 5.2). repeat experiments in which conditions were changed to attempt to minimise proteolysis did not increase the proportion of intact fusion protein.

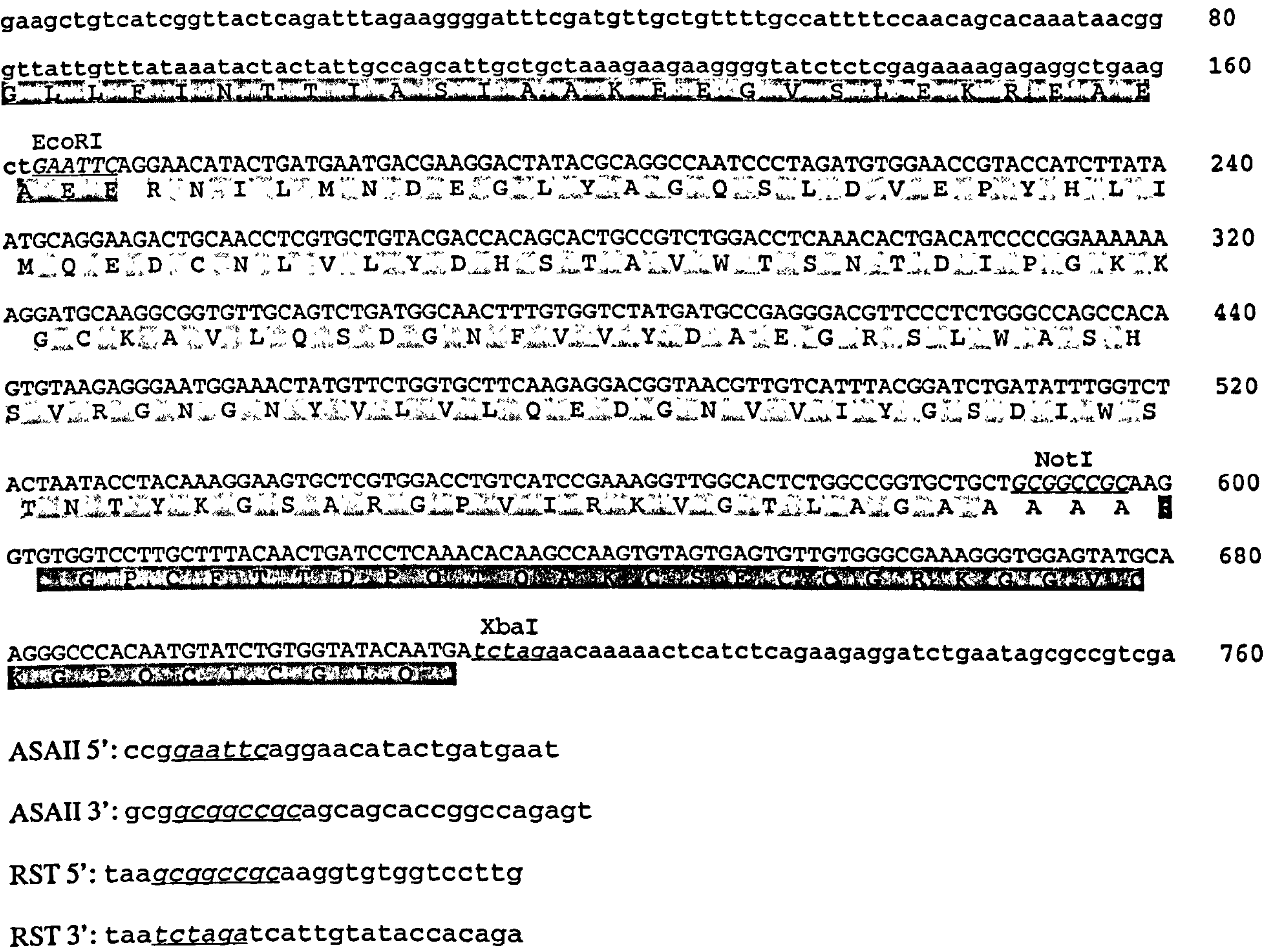


Figure 5.1 Sequence information for full-length ASaII₁₂₇-RST. Bases in uppercase denote recombinant protein sequence, grey highlighting shows translated vector sequence (α -factor), yellow highlighting indicates ASaII sequence, green highlighting shows RST sequence. Italic and underline indicates restriction sites. Primers used to amplify the coding sequence fragments used to assemble the construct are also shown.

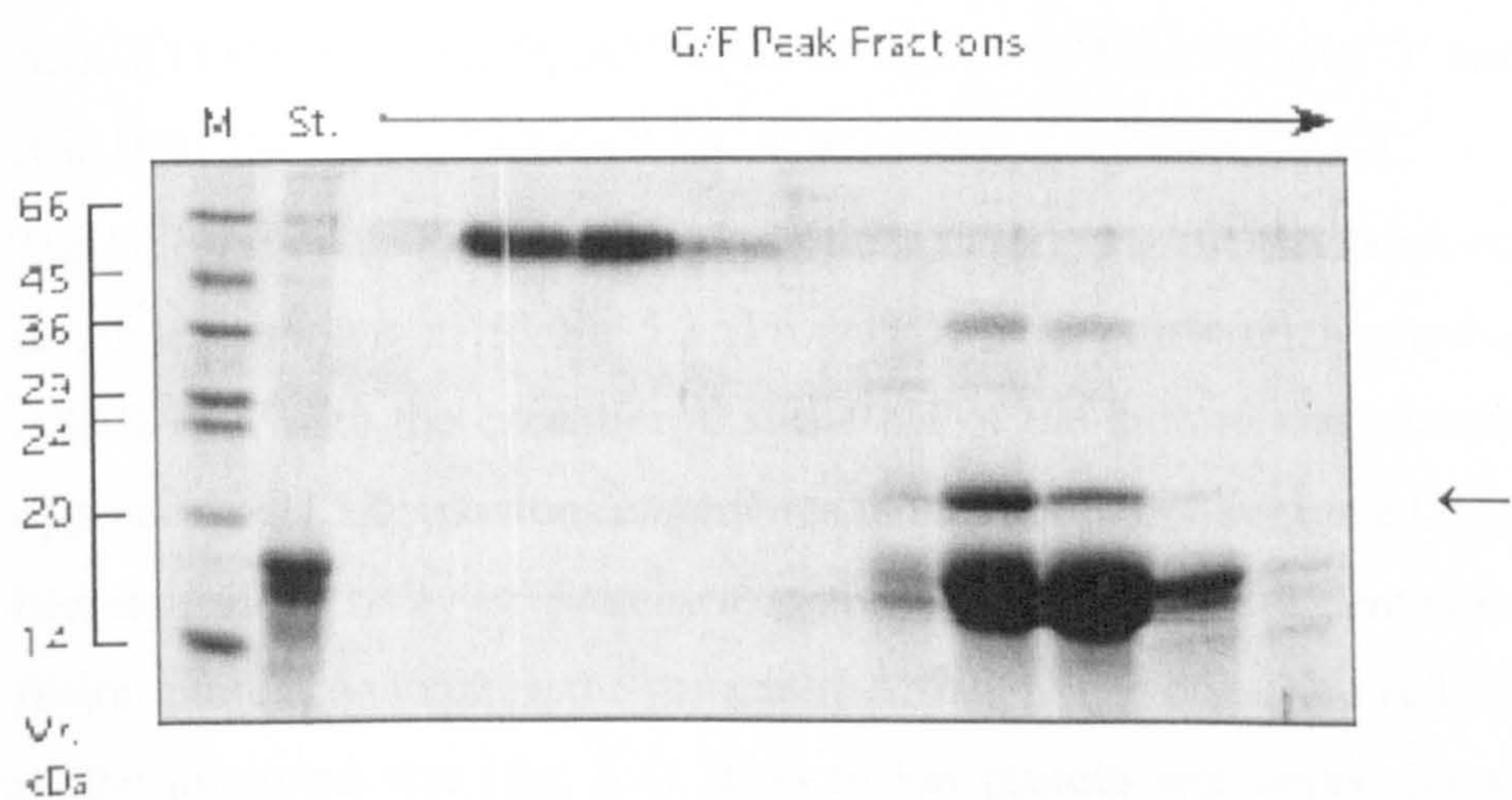


Figure 5.2 SDS-PAGE analysis of ASaII₁₂₇-RST after purification by gel filtration on Sephacryl S-200 column. Gel stained with Coomassie blue. Elution volumes of fractions increases in the direction of the arrow. The lane “St.” is a recombinant ASaII standard; arrow indicates undegraded fusion protein.

5.2.2 The ASaII₁₁₅-RST construct

A construct which encoded ASaII up to the predicted cleavage site (amino acids 1-115) was produced after MALDI-TOF revealed that cleavage of the ASaII C-terminus was leading to degradation of the full-length ASaII-RST fusion protein. To create this construct a PCR fragment was produced using the following oligonucleotides: ASaII(115)fwd: 5'- CCGAATTCAGGAACATACTGATGAAT-3' and ASaII(115)rev: 5'- TGCGGCCCGCAGCAGGTCCACGAGCACT.

The PCR product was amplified from the existing ASaIIpGAPzαA construct. Sequence details are shown in figure 5.3. Production and purification were carried out as for ASaII-RST with the exception that the pH of the culture was maintained at pH 4.0 as opposed to pH 5.0, previous experience in the laboratory suggested that lowering pH may decrease the activity of proteases in the culture during fermentation. Analysis of this fusion protein containing the truncated ASaII polypeptide showed intact fusion protein of the predicted size (fig. 5.4). This fusion protein was tested for toxicity towards *A. pisum*.

5.2.3 Artificial Diet Bioassay of ASaII₁₁₅-RST against *A. pisum*

Artificial diet bioassays of ASaII₁₁₅-RST at a range of concentrations from 0.1 mg/ml to 1 mg/ml in diet demonstrated non-toxicity to *A. pisum* over 11 days. No significant acute or chronic toxicity was observed during the course of the assay (fig. 5.5). This result is unexpected on two counts: first, the ASaII part of the fusion has not shown the expected toxicity; and second, the RST part of the toxin has had no apparent effect. This could be explained by the inability of ASaII to bind to the insect gut or to cross into the insect haemolymph when attached to the RST subunit. Alternatively, transport may occur to the haemolymph but the RST subunit may not be toxic to this species of insect. However, a similar lack of toxicity has been observed in the lepidopteran species *Lacanobia oleracea*, the tomato moth (Dr. E. C. Fitches, personal communication). RST has been reported as a lepidopteran specific toxin and FP4 (a fusion protein consisting of GNA-RST) has been demonstrated to show toxicity against *L. oleracea* as well as the

hemipteran pest *N. lugens*. This makes it likely that the problem of non-toxicity lies in the inability of RST to be delivered to the insect haemolymph by ASAIL.



ASaII 5': ccggaattcaggaacatactgatgaat
ASaII 3': tgcggccgcagcaggtccacgagcact
RST 5': taagcggccgcaaggtgtggtccttg
RST 3': taatctagatcattgtataccacaga

Figure 5.3 Sequence information for truncated ASaII₁₁₅-RST. Bases in uppercase denote recombinant protein sequence, grey highlighting shows translated vector sequence (α -factor), yellow highlighting indicates ASaII sequence, green highlighting shows RST sequence. Italic and underline indicates restriction sites. Primers used to amplify the coding sequence fragments used to assemble the construct are also shown.

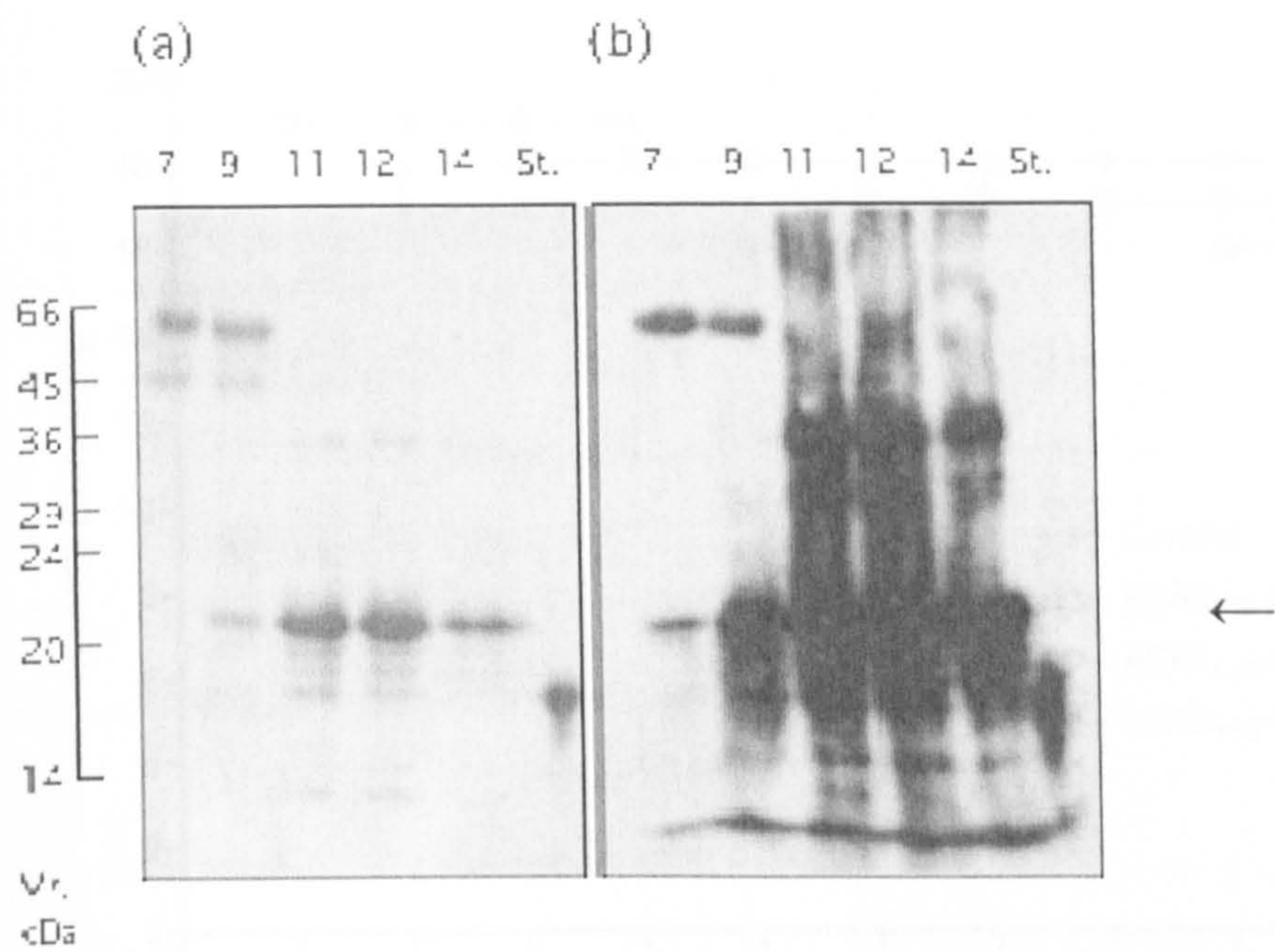


Figure 5.4 Analysis of ASAlI₁₁₅-RST by (a) SDS-PAGE (gel stained with Coomassie blue) and (b) Western-blot using anti-ASA antibody. Protein fractions were analysed after purification by gel filtration on Sephacryl S-200 column. Elution volumes of fractions increases with numbering. The lane “St.” is a recombinant ASAlI standard; arrow indicates undegraded fusion protein.

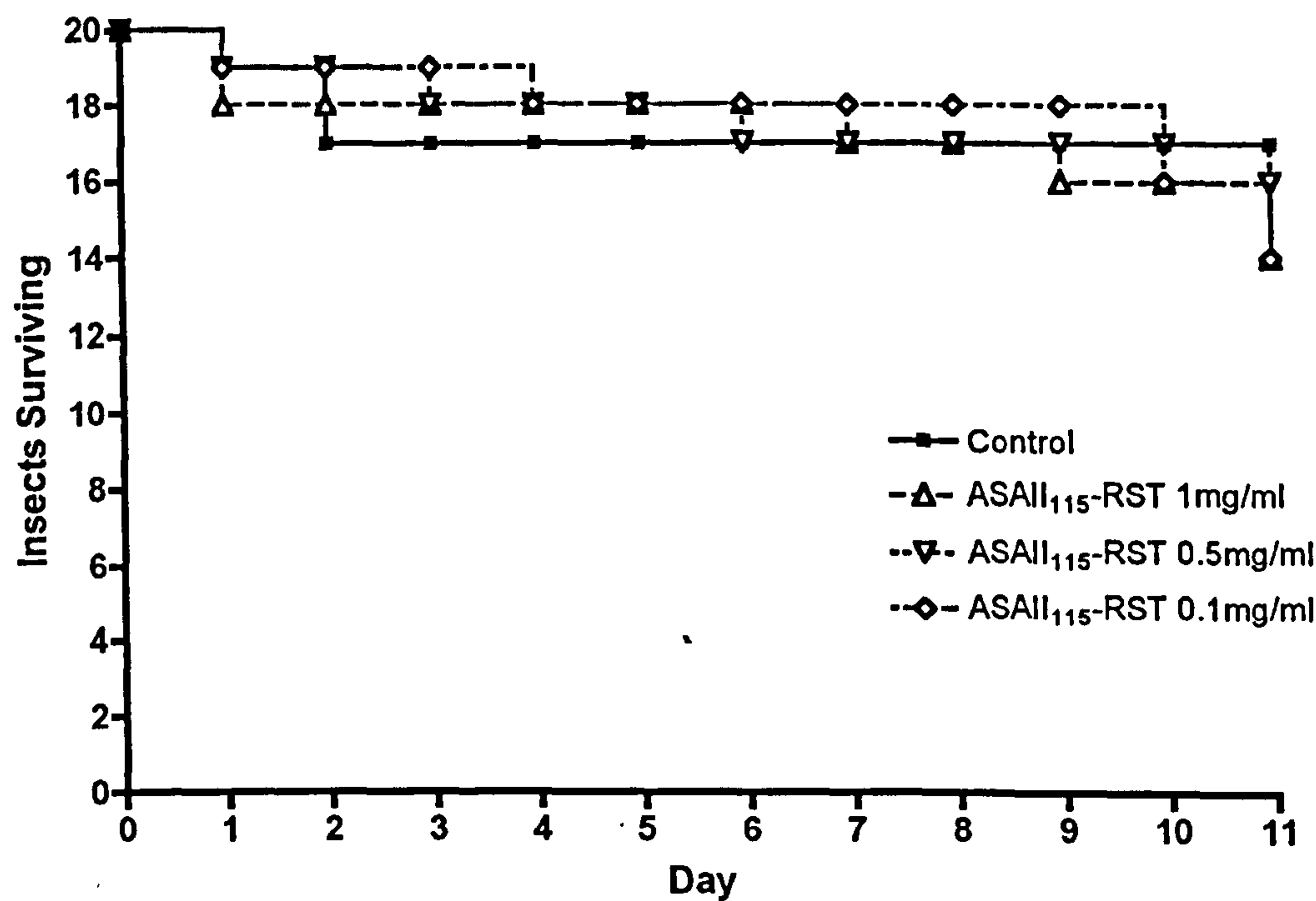


Figure 5.5 Bioassay of ASAIL₁₁₅-RST fusion protein against pea aphid (*A. pisum*).

5.3 The ASAlI-Avidin fusion protein

Cloning of chicken avidin and preparation of the ASAlI-Avidin expression construct were carried out by Dr. D. P. Bown.

5.3.1 Extraction of genomic DNA from chicken (*Gallus gallus*) embryo

200 mg of tissue from a five day old chicken embryo was homogenized in Tri reagent (Sigma; buffered guanidinium thiocyanate solution) and extracted with chloroform. After centrifugation to remove insoluble debris, DNA was precipitated from the aqueous phase by addition of EtOH. The DNA was pelleted by centrifugation, washed, and dried. The pellets were resuspended in 1000 μ l 8 mM NaOH overnight at 4 °C. To dissolve gelatinous material, the solution was heated to 60 °C for 5 minutes, and then cooled to 4 °C and neutralised by adding 0.1 M HEPES. The resulting DNA solution was analysed by spectrophotometry; it had an A_{260}/A_{280} ratio of 1.7, and a total yield of DNA of approximately 900 μ g.

5.3.2 Isolation of avidin coding sequence from chicken genomic DNA and assembly of avidin expression construct

Since attempts at amplification of an avidin cDNA by RT-PCR was unsuccessful, genomic DNA as described above was used as a template to amplify an avidin coding sequence. The avidin gene from *G. gallus* contains 4 exons, with the majority of the coding sequence of the mature polypeptide contained in exons 2 and 3 (see fig. 5.6). Exons 2 and 3 were amplified separately using primers Ex12f and Ex23r (exon 2; see fig. 5.7) and Ex3f and Ex3r (exon 3; see fig. 5.7) respectively. After isolation by gel electrophoresis, band excision and DNA purification, the exons were cloned separately into pCR2.1 using the TOPO-TA method, and selected clones were checked by DNA sequencing. An expression construct for the complete avidin coding sequence was then assembled in a two-step process (see fig. 5.8). In the first step, exons 2 and 3 were assembled together by restricting the pCR2.1 clones with PstI and ScaI (exon 2) or ScaI

and EcoRI (exon 3) and ligating the isolated DNA fragments into pUC18 which had been restricted with EcoRI and PstI. The ligation product was cloned, and selected clones were checked by restriction analysis and DNA sequencing. In the second step, the combined exons 2 and 3 were assembled with oligonucleotides containing the exon 4 sequence (Ex4f and Ex4r; see fig 5.7), and cloned directly into the yeast expression vector pGAPz α B. Exons 2 and 3 were excised from the pUC vector by restriction with MluI and PstI, and the isolated DNA fragment was ligated to pGAPz α B that had been restricted with PstI and Sall, and to annealed oligonucleotides Ex4f and Ex4r. The ligation mixture was cloned, and selected clones were checked by restriction digestion and by DNA sequencing. A clone containing the complete expression construct (pGAPz α B-avidin) was selected for further work; the predicted amino acid sequence of the translation product is shown in figure 5.8.

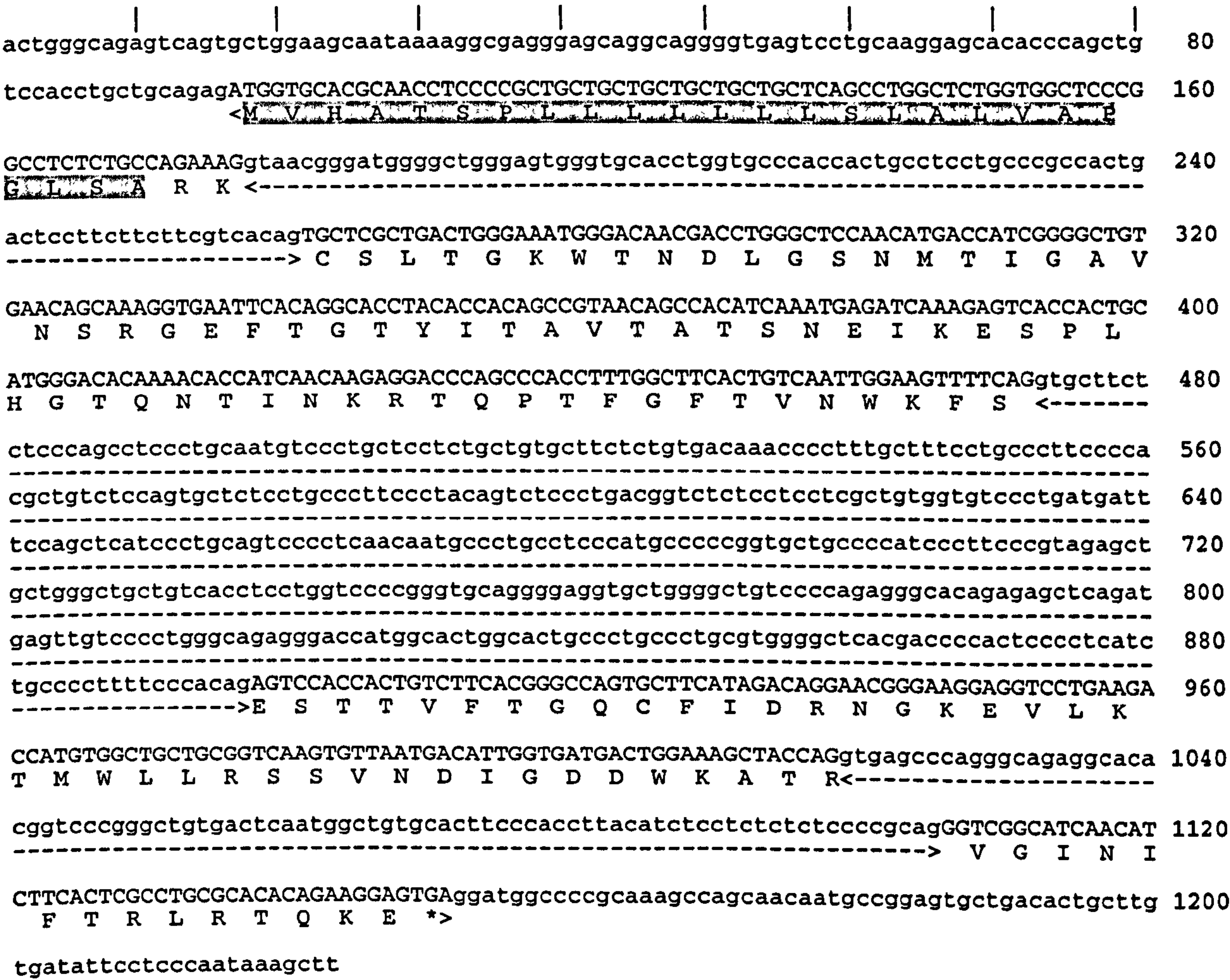


Figure 5.6 Avidin gene sequence (*Gallus gallus* AVD gene; AJ311647) showing intron/exon structure. Coding sequence is shown in capitals, non-coding in small characters. Introns are indicated by <--- --->. Shaded amino acid sequence indicates a signal peptide (predicted by SignalP v. 3.0).

SalI

Ex12f

5'

CGCGTCGACGCTAGAAATGCTCGCTGACTGGGAAATGG

3'

ScaI

Ex23r

5'

CGCGAGTACTTTCTGAAAACCTCCAATTGACGGTG

3'

ScaI

Ex3f

5'

CGCGAGTACTACTGTCTTCACGGGCCAGTGC

3'

MluI

Ex3r

5'

CGCGACGCGTAGCTTTCCAGTCATCACC

3'

Ex4f

5'

CGCGTGTGTTGGTATCAACATCTTCACTAGATTGAGAACTCAAAGGAAA

3'

Ex4r

3'

ACAACCATAGTTGTAGAAGTGATCTAACTCTTGAGTTTTCCTTTAGCT

5'

Figure 5.7 Primers and oligonucleotides used to assemble avidin coding sequence. Shaded bases show non-coding extensions.

(A)

(B)

Avidin expression construct

Yeast α -mating factor preprosequence

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLECFDVAVLFFSNSTN

NGLLFINTTIIASIAAKEEGVSLEKREAEA;AEFALRVDA

RKCSLTCKWTNDLGSNMTIGAVNSKGEFTGTYTTAVTATSNEIKESPLHGTONTINKRTO

PTFCFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDICDDWKATRVGINIFT

RLRTOKE

IDHHHHHH

(His)₆ tag

Avidin mature polypeptide

Sequence from expression vector

avidin mature polypeptide

Sequence from expression vector

Figure 5.8 (A) Schematic diagram of avidin expression construct. Site in brackets is lost in the complete construct. (B) Predicted polypeptide produced by avidin expression construct.

- -

The avidin expression construct was checked by transformation into *P. pastoris* strain X-33. Transformed yeast colonies were selected by plating on zeocin-containing media (100 µg/ml). Selected transformants were grown in small-scale cultures for 3 days, and culture supernatants were screened by immuno-dot blot assay, probing with anti-avidin antibodies. The transformant giving the strongest signal was selected for further study. Western blot analysis of culture supernatant showed the presence of a polypeptide of molecular weight approximately 20,000 Da when probing with anti-avidin antibodies. The polypeptide was approximately 2,000 Da larger than “natural” avidin, as predicted by the extra amino acid sequence present in the sequence encoded by the expression construct (17 residues, N-terminal extension plus C-terminal (his)₆ tag – see fig. 5.8).

5.3.3 Preparation of ASaII-Avidin expression construct

The pGAPzαB-avidin expression construct was designed with extra coding sequence at its N-terminus, containing restriction sites to allow the insertion of an extra coding fragment (fig. 5.9). An ASaII fragment, corresponding to amino acid residues 1-115, and containing *Pst* I and *Sal* I restriction sites compatible with those in the pGAPzαB-avidin was produced by PCR amplification from the original pGAPzαA-ASaII expression construct using the following oligonucleotides:

GAPASAfwd: 5' CGCGCTGCAGAAAGGAACATACTGATGAATGACG 3'

ASAAvrev: 5' GCGCGTCGACAGGTCCACGAGCACTTC 3'

Restriction sites are underlined. The PCR product was purified by agarose gel electrophoresis, and DNA was purified from the excised band. The fragment was restricted with *Pst* I and *Sal* I, and was ligated to pGAPzαB-avidin that had been restricted with the same enzymes. The ligation mixture was transformed into *E. coli*, and plated on zeocin-containing media. Putative transformants were checked by restriction digestion of plasmid DNA extracted from the colonies, and clones containing the correct construct were checked by DNA sequencing to obtain the expected pGAPzαB-ASaII-Avidin expression construct (fig. 5.9). The sequence of the predicted polypeptide product of this expression construct is shown in figure 5.10.

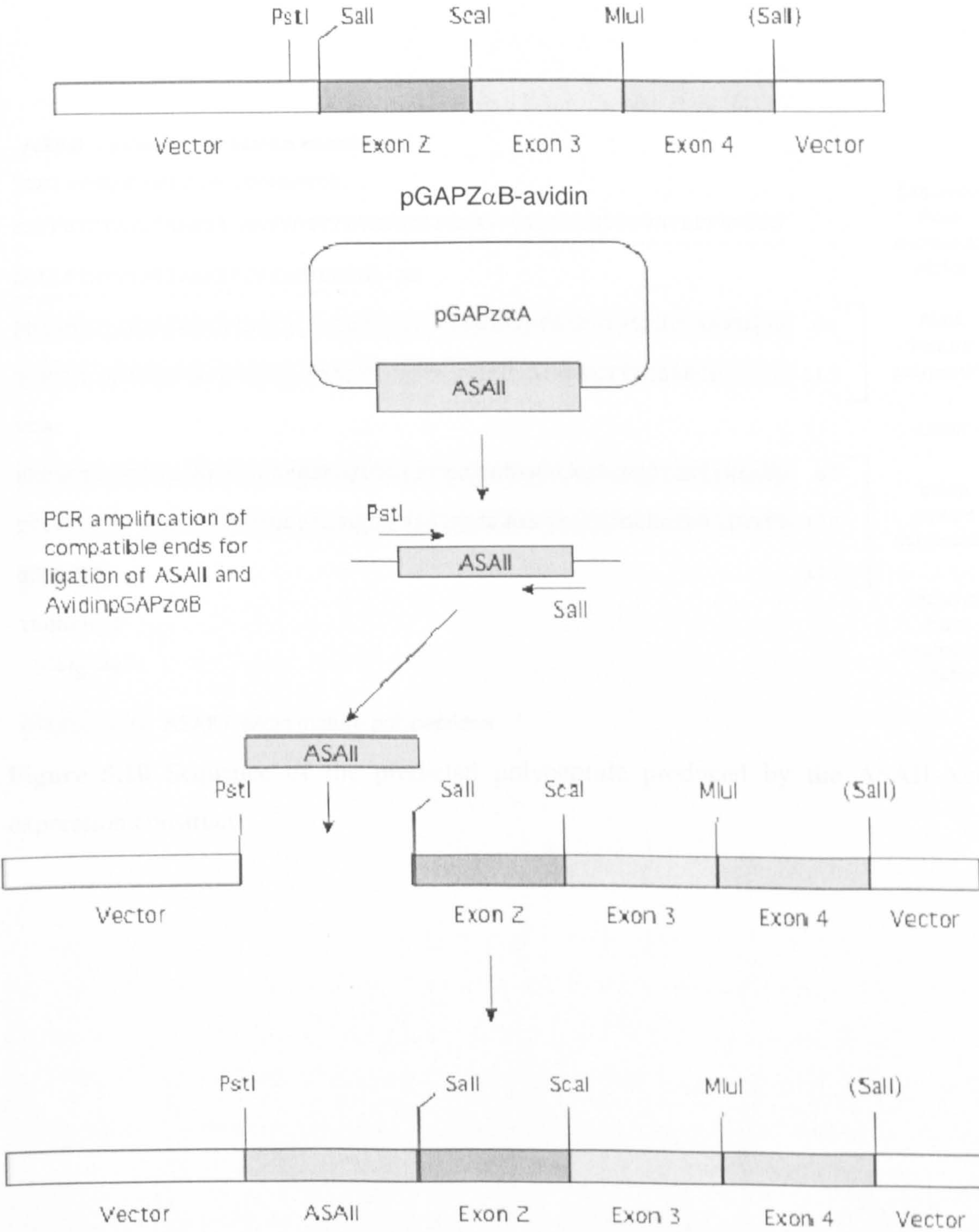


Figure 5.9 Schematic diagram showing the assembly of the ASaII-Avidin expression construct.

ASaII - avidin expression construct

Yeast a-mating factor preprosequence			
MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLECDFDVAVLPFSNSTN			Sequence from expression vector
NGLLFINTTIAASIAAKEEGVSLEKREAEA AE			
RNILMNDEGLYAGQSLDVEPYHLIMQEDCNLVLYDHSTAVWTSNTDIPGKKGCKAVLQSD	60	ASaI mature polypeptide	
GNFVVYDAEGRSLWASHSVRGNGNYVLVLQEDGNVVIYGSDIWSTNTYKGSARGP	115		
VDA			Linker
RKCSLTGKWTNDLGSNMTIGAVNSKGFTGTYYTTAVTATSNEIKESPLHGTQNTINKRTQ	60	avidin mature polypeptide	
PTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFT	120		
RLRTQKE	127		
IDHHHHHH			Sequence from expression vector
(His) ₆ tag			
ASaI / avidin mature polypeptides			

Figure 5.10 Sequence of the predicted polypeptide produced by the ASaII-Avidin expression construct.

5.3.4 Production and Purification of ASaII-Avidin

ASaII-Avidin was produced in *P. pastoris* (protease-deficient SMD1168 strain) as described for recombinant garlic lectins in Chapter 3. The validated clone expressing the fusion protein was grown in a 5 litre culture using a laboratory scale fermenter.

Initial purification of ASaII-Avidin by hydrophobic interaction chromatography using a phenyl-Sepharose column was not possible, as ASaII-Avidin failed to bind to the matrix, as was seen in the case of recombinant ASaII (results not shown). In line with the techniques used for recombinant lectins, cation-exchange chromatography using S-Sepharose matrix was used as an alternative. As for the purification of recombinant lectins by this method, a 10-fold dilution of the harvested culture supernatant was required to reduce the concentration of sodium chloride to a level which allowed proteins to elute across a 0-0.5 M NaCl gradient. The culture supernatant was diluted 10-fold into 50 mM sodium phosphate buffer (pH 4.0) and purification was carried out in 5 litre batches. Analysis on SDS-PAGE showed that ASaII-Avidin was present in the peak fractions eluted from the column by a 0-0.5 M [NaCl] gradient (fig. 5.11(a)), but analysis of the column flow-through showed that binding to S-Sepharose was not efficient (fig. 5.11(b)) with large amounts of recombinant protein in the flow-through volume.

As ASaII-Avidin was expressed in the protease-deficient SMD1168 strain of *P. pastoris*, the (his)₆ tag should be retained. This allowed affinity chromatography using metal affinity chromatography (Ni²⁺ matrix) to be carried out, which did not require dilution of initial culture supernatant. Nickel affinity chromatography was carried out initially in 50 mM sodium phosphate buffer at pH 7.4. 300 mM sodium chloride was used in the buffers as oppose to the recommended concentration of 500 mM to allow for the presence of sodium chloride in the culture supernatant. Loading, washing and elution steps were all carried out at 4 °C. When the pH of the buffered culture supernatant was raised to 7.4 a significant amount of precipitate formed, this precipitate was not the ASaII-Avidin fusion protein, which was still present in the filtered solution. However, the precipitation caused problems when running the metal affinity columns due to clogging. After washing

the affinity column, purified fusion protein was eluted using 300 mM imidazole in column buffer. Typically elution volumes for affinity chromatography are less than 5 ml, resulting in very high concentrations of protein in the elution solution. When dialysed against ammonium bicarbonate buffer to remove imidazole, further precipitation of the sample occurred, and in this instance it was recombinant protein that precipitated. To overcome this problem the eluted sample was diluted approximately 10-fold with distilled water prior to dialysis. At this lowered protein concentration no precipitation occurred. For subsequent experiments it was necessary to concentrate the recombinant ASaII-Avidin using a centrifugal concentrator, during which no precipitation of the protein occurred.

Since no precipitation of material from culture supernatant occurred during ion-exchange chromatography at pH 4.0, whereas precipitation occurred at pH 7.4 during metal affinity chromatography, affinity chromatography was attempted at pH 4.0 in 100 mM sodium acetate buffer. It was observed that increasing the pH as far as pH 5.0 initiated some precipitation in the culture supernatant, so pH 4.0 was a suitable pH to avoid this precipitation during the loading of the sample onto Ni²⁺ resin. Elution of the protein still had to be carried out at pH 7.4 because at pH 4.0 protein would not elute from the matrix (result not shown). This increase in pH during elution did not result in precipitation in the elution solution, provided that the column was washed to remove non-bound material at pH 4.0.

5.3.5 SDS-PAGE analysis of recombinant ASaII-Avidin produced in *P. pastoris* SMD1168

Analysis of ASaII-Avidin purified by ion-exchange chromatography on SDS-PAGE produced a profile of major bands at approximately 18 kDa, 25 kDa and approximately 60 kDa (strongest band; see fig. 5.11). However, analysis of fractions and flow-through by Western blotting showed that incomplete binding to the ion-exchange column was occurring, with immunoreactive material present in the flow-through fraction, and that

the major stained band did not correspond to the major immunoreactive species (fig. 5.11(b)). This method of purification was not pursued further.

Purification of ASaII-Avidin by Ni^{2+} affinity chromatography gave a product eluted from the column by imidazole that contained a major band of approximately 36 kDa (fig. 5.12 (b)) that corresponded to a similar sized immunoreactive band on a Western-blot carried out with anti-avidin antibody (fig. 5.12(c)) and anti-myc tag antibody (fig. 5.12(d)). This band represents intact ASaII-Avidin; the immunoreactivity with anti-myc tag antibody demonstrates that no C-terminal cleavage of the avidin subunit is occurring in SMD1168. Further bands are seen at approximately 20 kDa and 18 kDa. The 20 kDa band is strongly immunoreactive with anti-avidin antibody suggesting it is cleaved avidin. Avidin usually runs as a 16 kDa monomer, suggesting that the 20 kDa band contains extra amino acids than is seen for avidin monomer. This may be due to the cleavage of ASaII at its C-terminus as described in chapter 3. If this were to be occurring then an immunoreactive band containing the C-terminus of ASaII plus the linker region and the avidin subunit of ASaII-Avidin would be produced, the molecular weight of which would be approximately 20 kDa. This band is also immunoreactive with the C-terminal anti-myc tag antibody (fig. 5.12(d)).

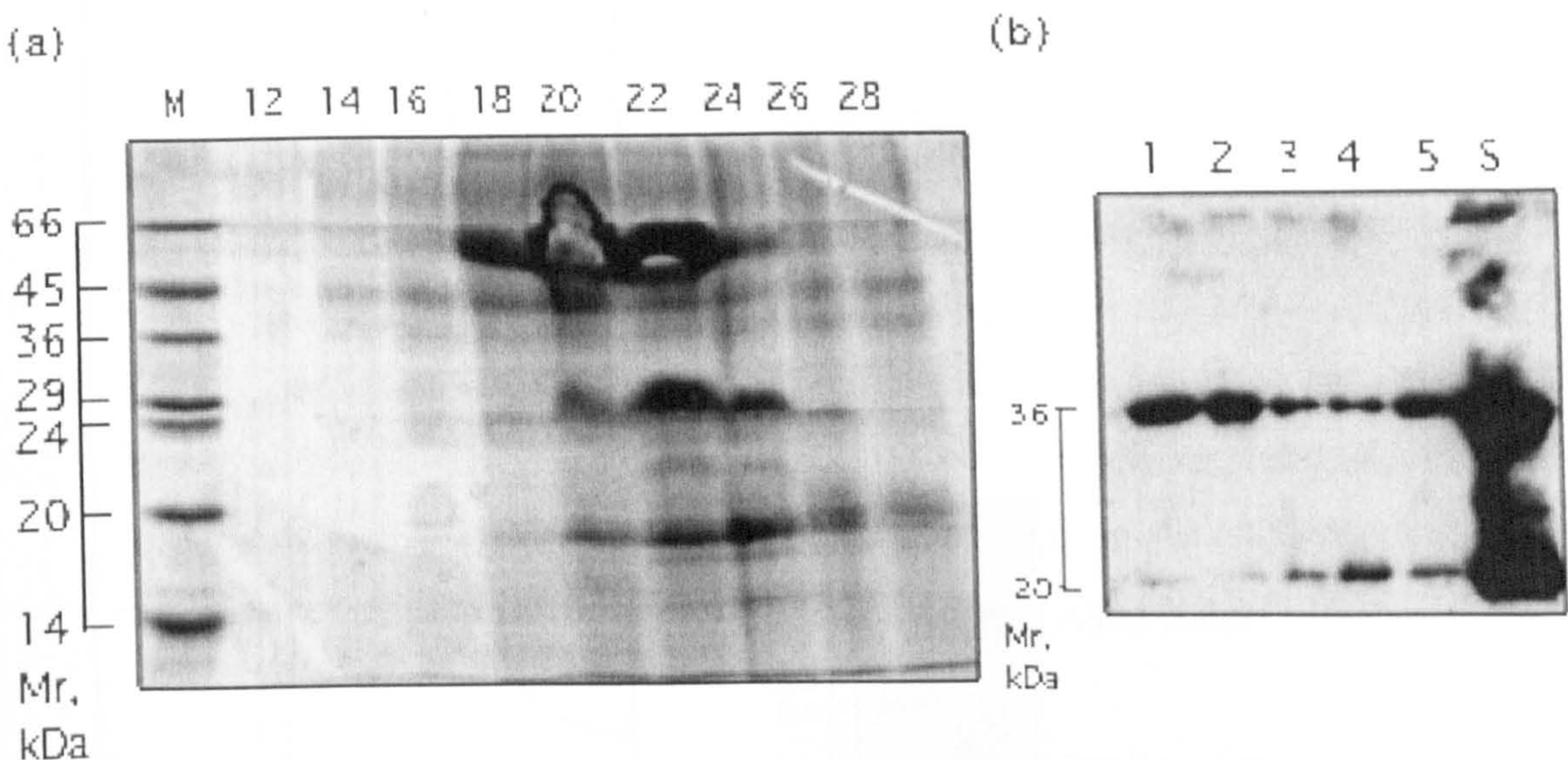


Figure 5.11 Purification of ASaII-Avidin from yeast culture supernatant by ion-exchange chromatography on S-Sepharose at pH 4.0. (a) SDS-PAGE of peak fractions taken from peak fractions (300 μ l of the fraction acetone precipitated prior to loading on gel). (b) Western blot analysis of flow-through from various ion-exchange purifications. 1-5 indicate different purification runs, S = starting sample. Detected with anti-Avidin antibodies.

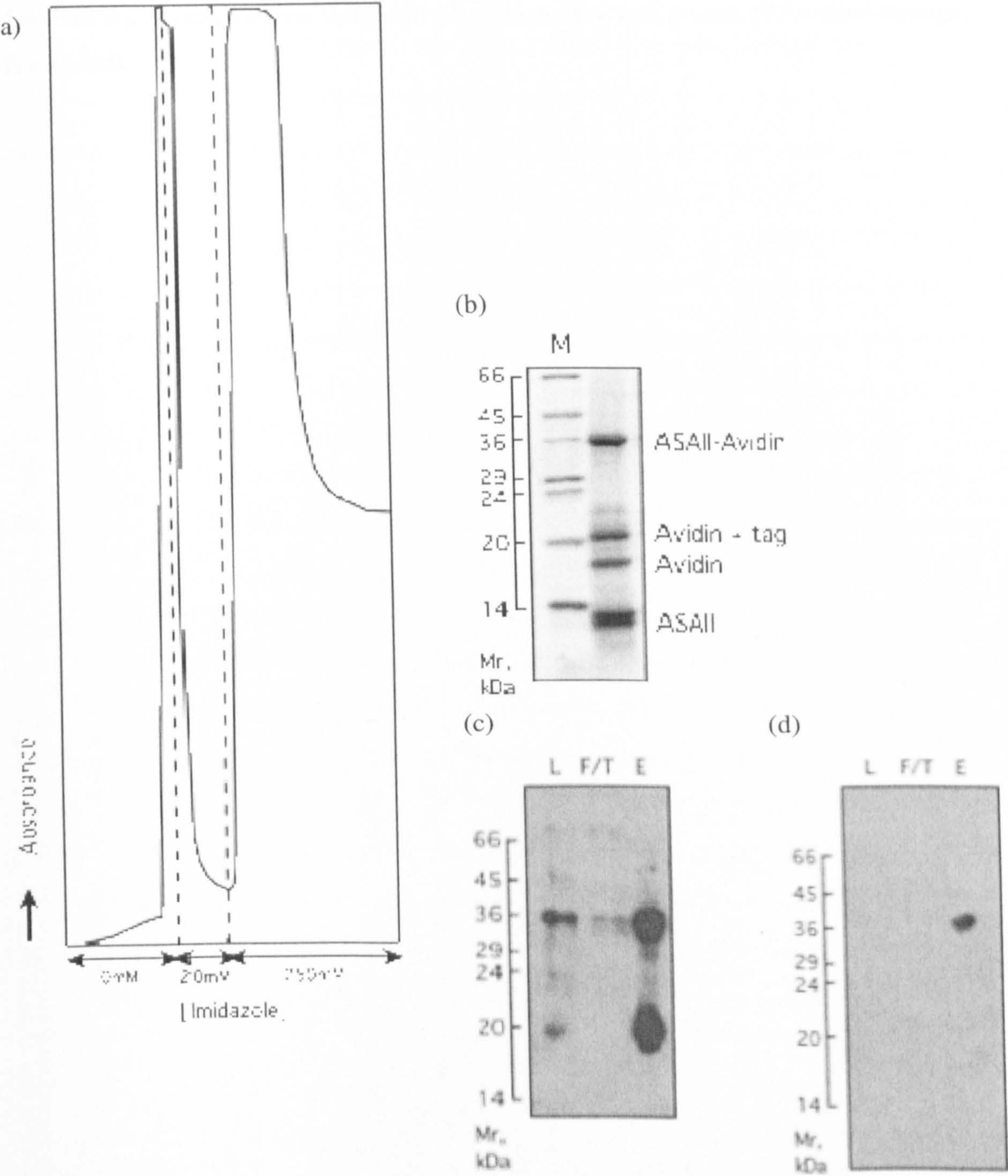


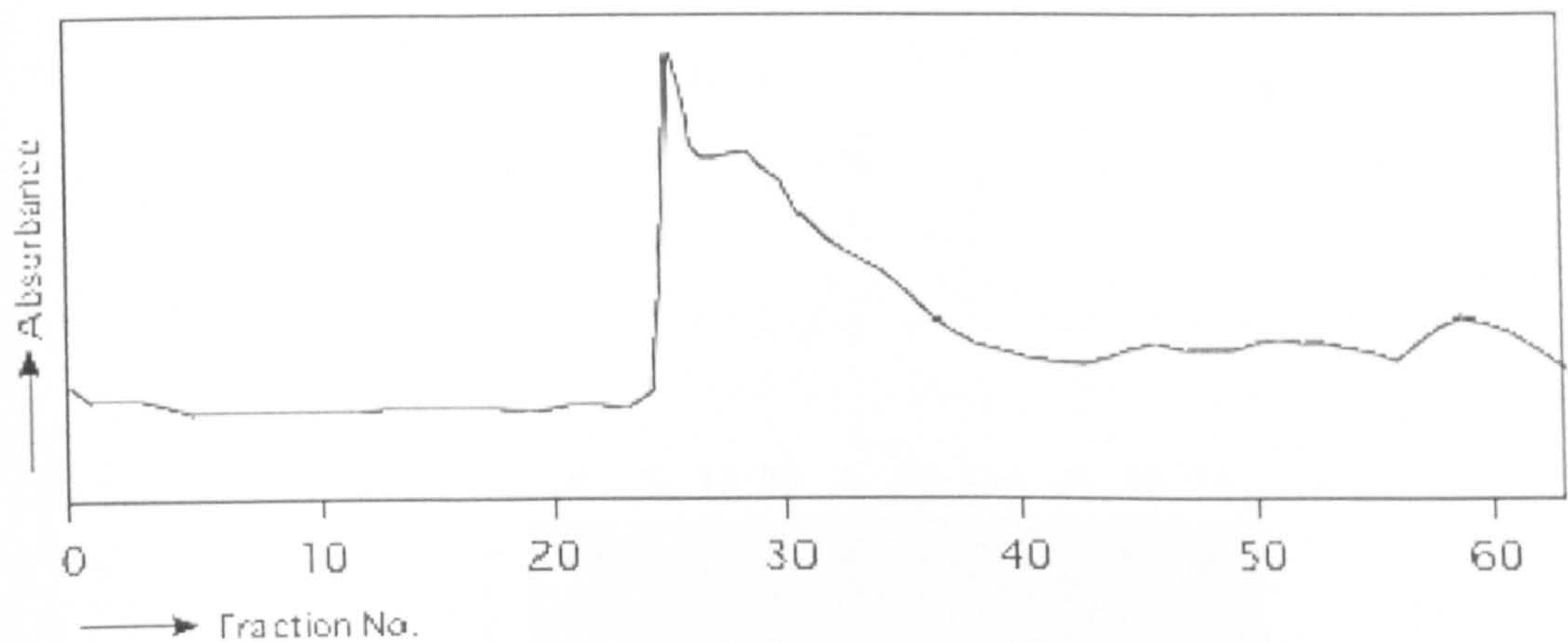
Figure 5.12 Purification of ASAll-Avidin expressed in *P. pastoris* strain SMD1168. (a) Purification of ASAll-Avidin from SMD1168. Elution profile of Ni^{2+} affinity column; (b) SDS-PAGE analysis of fraction eluted from column in 250 mM imidazole (after gel-filtration clean up and centrifugal concentration of peak); (c) Western-blot analysis of fractions eluted from column using anti-Avidin antibody; (d) Western-blot analysis of

fractions eluted from column using anti-Myc tag antibody (L = load, F/T = flow-through, E = eluted).

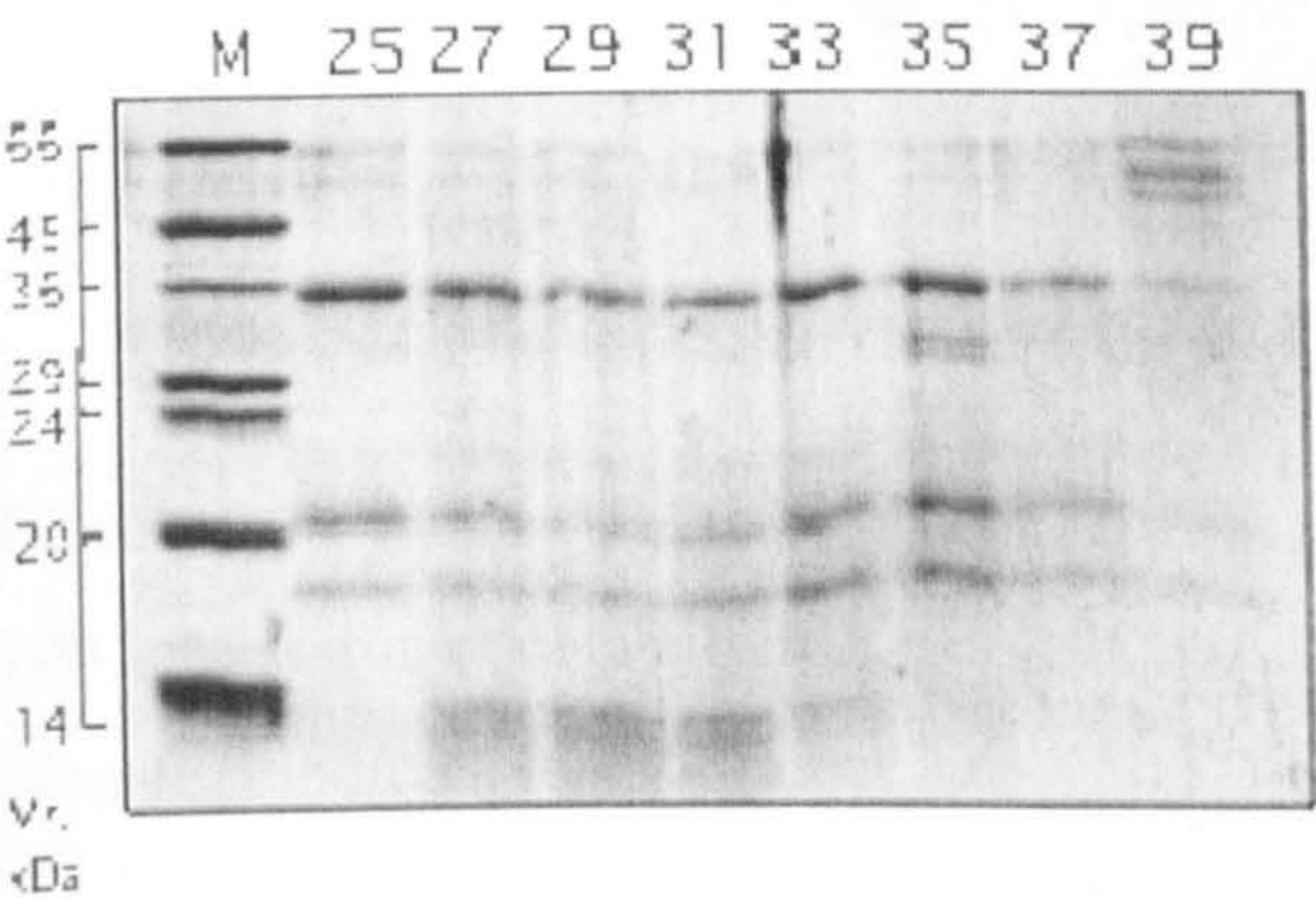
5.3.6 ASaII-Avidin produced in *P. pastoris* strain SMD1168 shows variability in proteolytic degradation

Maintaining intact ASaII-Avidin throughout the purification procedure was problematic and as was seen with recombinant garlic lectins expressed in *P. pastoris*, levels of proteolytic degradation were variable. After several attempts, intact ASaII-Avidin was produced with a minimal amount of degraded products. In the ion-exchange purification protocol, lengthy loading periods (upto 5 days in some instances) at 4 °C and further gel-filtration “clean-up” chromatography led to degradation of ASaII-Avidin into its constituent subunits (fig. 5.13). Purification by Ni²⁺ affinity in sodium phosphate buffer pH 7.4 led to precipitation of salt in the buffer and later of protein during dialysis. Intact recombinant protein production was achieved by carrying out purification on Ni²⁺ affinity matrix at pH 4.0 in sodium acetate buffer and elution at pH 7.4 in loading buffer containing 300 mM imidazole, elution profiles for this purification were similar to those seen in the sodium phosphate strategy. Critical to the purification of intact recombinant protein was immediate dilution of the elution solution which was then dialysed against water immediately at 4 °C. Immediately following dialysis the sample was centrifugally concentrated at 4 °C, a sample taken for analysis by SDS-PAGE and BCA assay and the remainder immediately transferred to –20 °C for storage. Proteolytic degradation did not occur during –20 °C storage and the concentrated sample could be thawed for “clean-up” gel-filtration chromatography. The intact protein produced by this procedure is shown (fig. 5.14).

(a)



(b)



(c)

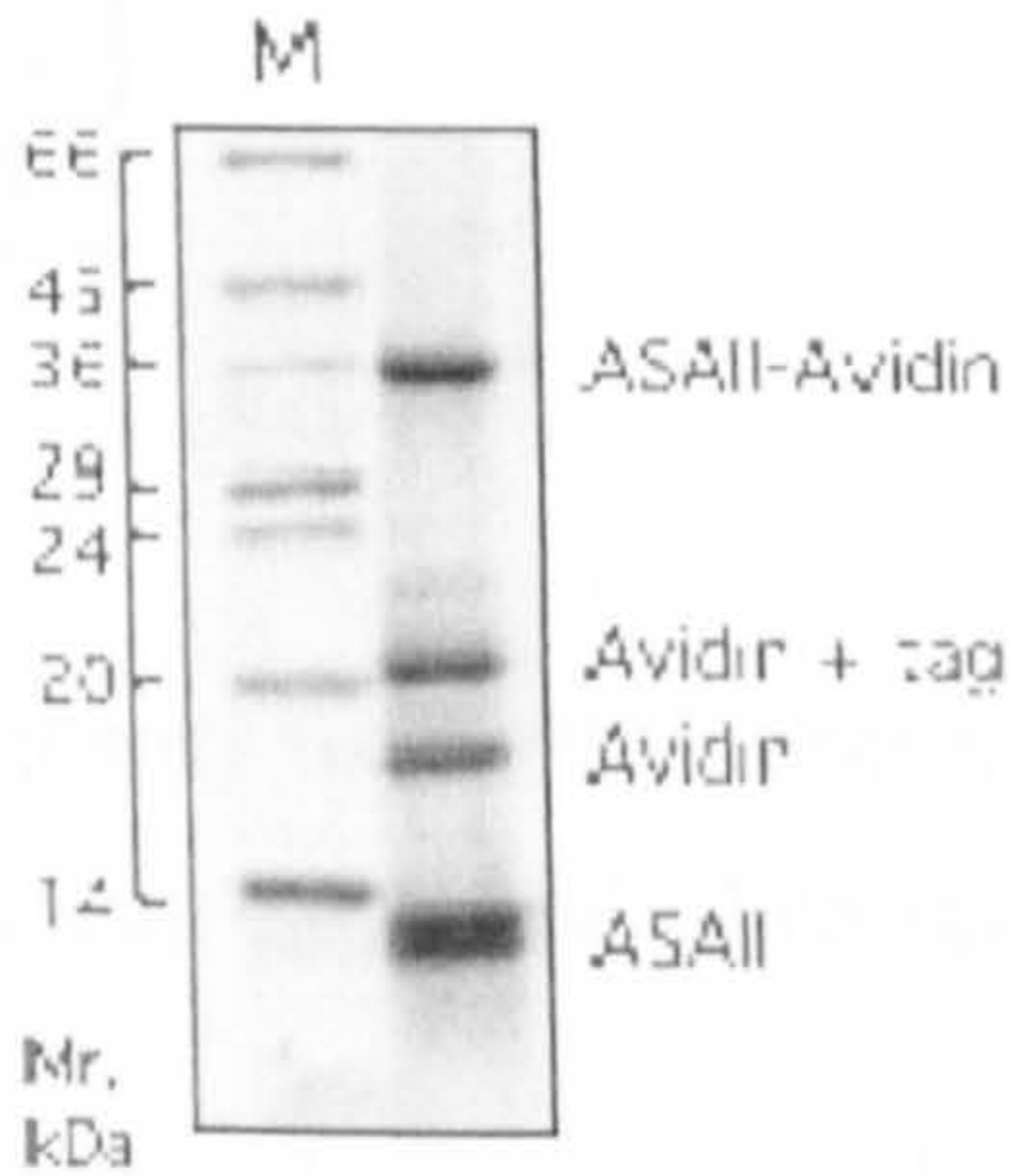


Figure 5.13 “Clean-up” purification profile of ASaII-Avidin. Gel-filtration was carried out on a column of Sephacryl S-200. (b) SDS-PAGE of samples from ASaII-Avidin expressed as a recombinant protein in *P. pastoris* strain SMD1168. (c) Pooled peak fractions of ASaII-Avidin after centrifugal concentration (MWCO 10 kDa).

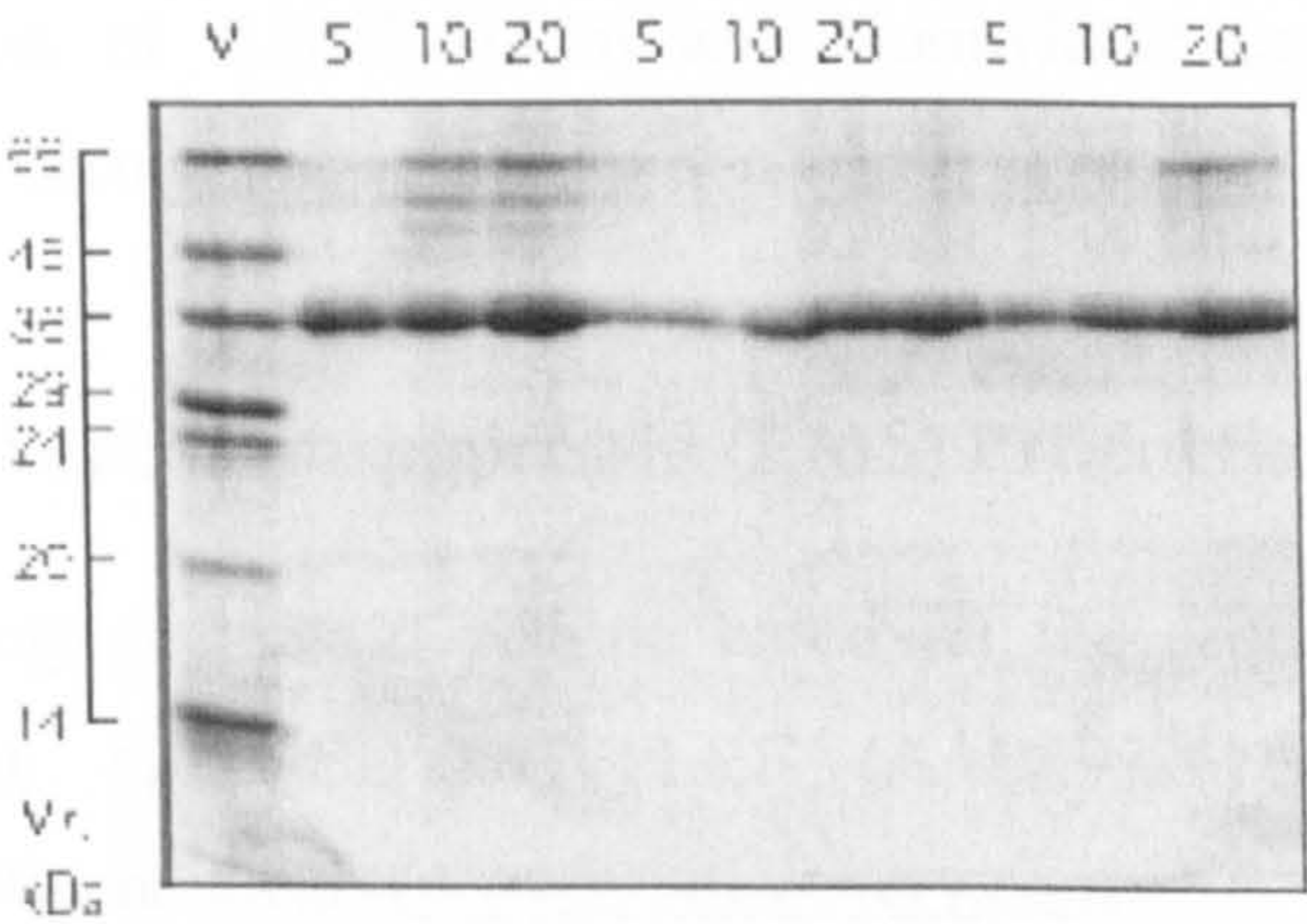


Figure 5.14 Intact ASaII-Avidin following affinity chromatography using Ni²⁺ resin in 100 mM sodium acetate buffer pH 4.0 (elution at pH 7.0) and ‘clean up’ by gel-filtration. Numbers indicate load volumes (μl).

5.4 ASaII-Avidin:Biotin-Leucomyosuppressin

The production of recombinant fusion proteins for analysis of small potentially insecticidal peptides by cloning based strategies is a laborious task for screening a large number of peptides. The availability of synthetically-produced peptides with biotin tags enables ASaII-Avidin to be utilized as a tool for the transport of small peptides without the need for complex cloning procedures. The strong binding between biotin and avidin allows the conjugation of ASaII-Avidin and biotinylated peptides to proceed to completion in a simple incubation.

5.4.1 Biotinylated Leucomyosuppressin (LMS) Production

A biotinylated leucomyosuppressin peptide based on the sequence from *Diploptera punctata* (cockroach), biotin-QDVDHVFLRF (AAB39926) was synthesized by Cambridge Research Biochemicals (www.crbdiscovery.com).

5.4.2 Conjugation of ASaII-Avidin and Biotin-LMS

Purified ASaII-Avidin was typically lyophilized and later resuspended in PBS for conjugation with biotin-LMS. The conjugation was carried out in the presence of a large molar excess of LMS; 2.5 mg (approximately 70 nmoles) of ASaII-Avidin was resuspended in 1 ml of PBS and 0.5 mg (approximately 300 nmoles) of LMS resuspended in 0.5 ml PBS. The two samples were then mixed and incubated overnight at 4 °C with rotation. Verification that the conjugation had occurred was carried out by initial separation by gel-filtration chromatography to separate conjugated ASaII-Avidin-LMS from unconjugated ASaII-Avidin and LMS subunits (fig. 5.15(a)) and analysis of peak fractions by Western blotting using anti-avidin antibodies (fig. 5.15(b)) and an ELISA using anti-FMRamide antibodies (fig. 5.15(c)).

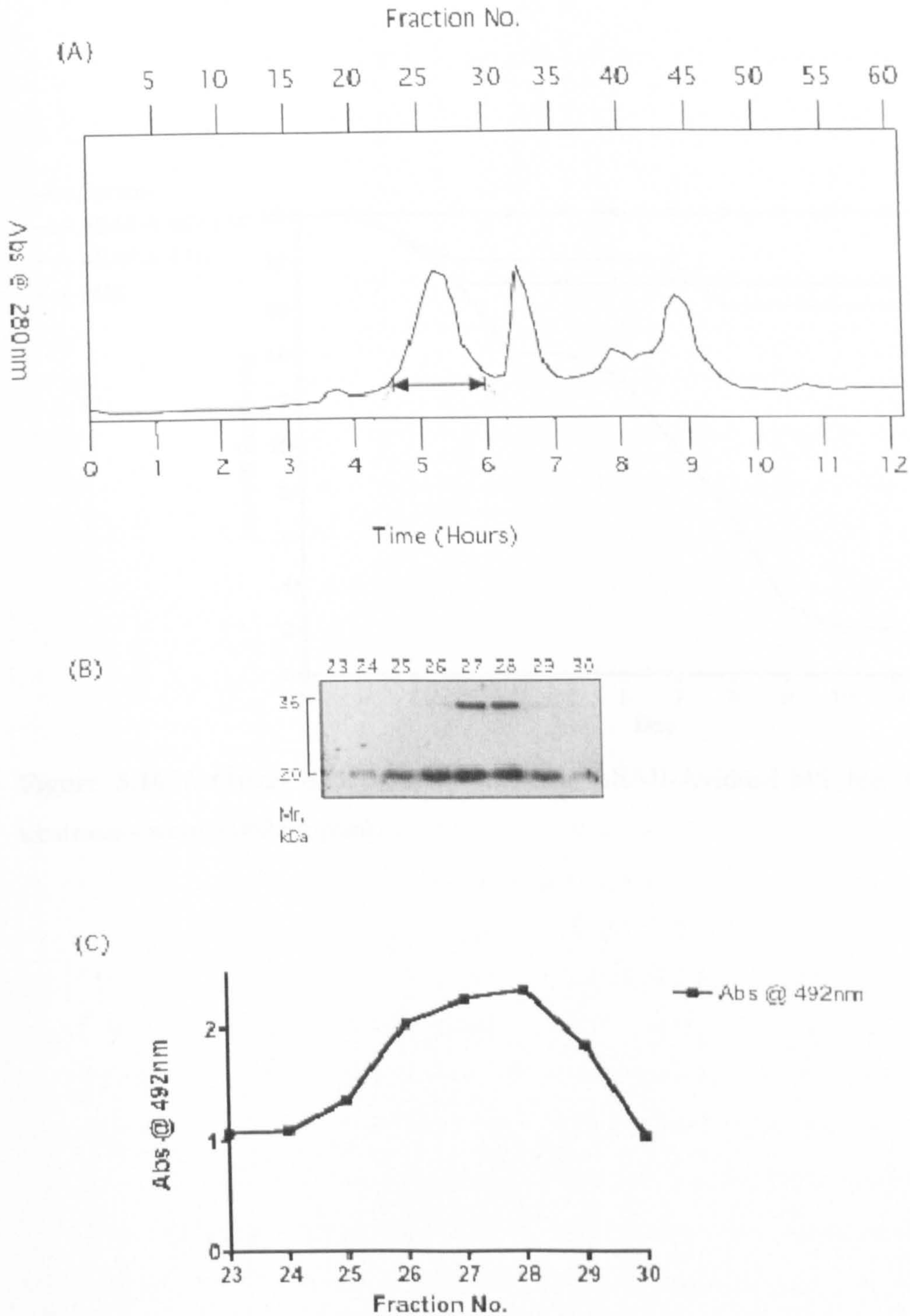


Figure 5.15 (A) Purification of ASAIL-Avidin-LMS by gel filtration chromatography; (B) Western blot analysis of gel filtration peak fractions with anti-avidin antibodies; (C) Analysis of gel filtration peak fractions by ELISA using anti-FMRF antibodies.

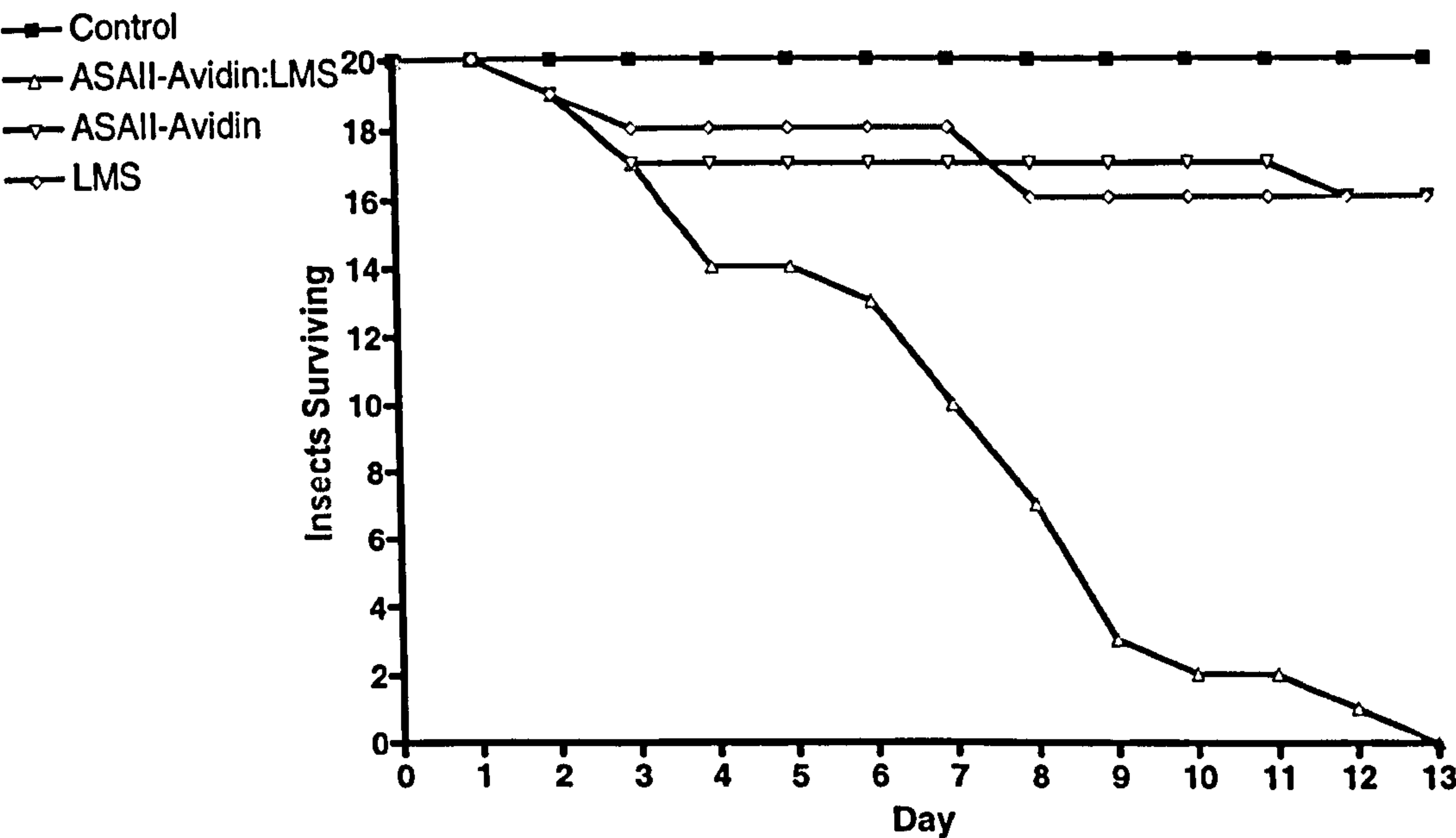


Figure 5.16 Artificial diet bioassay data for ASAIL-Avidin-LMS fed *A. pisum*. All treatments were 0.0625 mg/ml.

5.4.3 Artificial Diet Bioassays of ASaII-Avidin-LMS against *A. pisum*

Artificial diet bioassays were carried out against *A. pisum* to test for insecticidal activity of ASaII-Avidin-LMS. Control treatments for these assays had to be carefully designed, since avidin itself is known to have an insecticidal effect when fed. This effect is dependent on the biotin-binding activity of free avidin, since feeding avidin in the presence of biotin does not give an insecticidal effect. Since the ASaII-Avidin-LMS conjugate had been prepared with an excess amount of biotinylated LMS, all biotin binding sites on the avidin would be blocked. To provide a meaningful ASaII-Avidin control, it was therefore necessary to block the avidin binding sites with excess biotin. ASaII-Avidin was incubated with a 2 x molar excess of biotin and after 1 hour unbound biotin was dialysed out of the sample. This biotin-blocked ASaII-Avidin was used in the artificial diet feeding trials. An LMS only control was also included, as the ability of LMS to cross the insect gut is unknown.

As described previously, day 0 *A. pisum* nymphs were placed on artificial diets containing the respective diet treatments. Control survival was 100% over the course of the assay (13 days). Survival of aphids fed both ASaII-Avidin and LMS only was 85% at the end of the assay, with no significant no acute or chronic toxicity seen in either of these two treatments. Survival of ASaII-Avidin-LMS fed aphids was 80% at day 4, displaying a limited acute toxicity, 10% at day 10 and 0% at day 13. Significant chronic toxicity was observed between days 4 and 10 with limited acute effects from day 0 to 4. This type of toxicity is consistent with the expected effect of the LMS; the hormone would be expected to result in perturbation of normal gut peristaltic movement, and thus to affect feeding. A progressive inability of the aphids to feed properly explains the observed effects. The results of the feeding bioassay suggest that LMS has a fatal effect on *A. pisum*, but when orally administered this effect is only apparent when LMS is conjugated to a carrier protein, in this case ASaII-Avidin. Both ASaII-Avidin and LMS only do not display toxicity to *A. pisum*, validating the role of the carrier protein in delivering insecticidal peptides to produce a novel insecticidal action. A summary of the results is shown in figure 5.16.

5.4.4 *A. pisum* contains a gene encoding a putative LMS peptide hormone

LMS-like peptide hormones have not been identified in aphids and other hemipterans, and thus the toxicity of cockroach LMS to pea aphid demonstrated above may have little relevance to endogenous hormones in aphids. However, examination and analysis of the sequence data available in the partially complete *A. pisum* genomic database (Aphidbase; www.aphidbase.com), taken in conjunction with EST data for other aphids, shows that *A. pisum* does indeed contain a gene encoding a precursor for an LMS-like peptide hormone. The assembled sequence for this precursor is shown in figure 5.17. The predicted *A. pisum* LMS hormone has the sequence QDL DHVFLRF, differing only in a single conservative amino acid substitution at position 3 from the *D. punctata* sequence used to make the ASaII-Avidin-LMS conjugate. The conclusion that LMS administered through the ASaII-Avidin conjugate is having a direct effect on endogenous regulatory processes in the pea aphid is therefore justified.


```
ATGATTGTGTACCTATCTCGTCTCTATAATATATTATGTAACACGCGGTGCAGTGCGGAAATGTTTATAAACTCTATCGT 80
M I V Y L S R L Y N I L C N T R C S A E M F I N S I V
ATTGATTGTTGTAGAAAGAAAAATAATGTCTACGTACAGGACGATGCTGACTAGCGTGCTGGTGCTACTGGTCGTGGCCA 160
L I V V E R K I <M S T Y R T M L T S V L L V L L V L V A
GCGTTTGCGATTGCCGGATAAACAACCTCCCCACGCGGTGTATGCCAGGGCTGCTGGAGGACGCGCCGCCCAAGGTGCGA 240
S V C D C R I N N L P T R C M P G L L E D A P P K V R
GAAGCGTGTTTGACTTTGTCCACCATCCGCACCCTATCGAACGCCATCGAAACATTCATCCAAGACAAACAGTTCCCGAT 320
E A C L T L S T I R T L S N A I E T F I Q D K Q F P M
      V
GCCTATGCCTTACAGCAGAATGTCTGATGGATTGGTCGATACAGCGATGCAAAATGATAAACGCCAAGACCTGGACCACG 400
P M P Y S R M S D G L V D T A M Q N D K R Q D L D H
TATTCCTCAGATTCGGTCGCAGAAGACGTTAA
V F L R E G R R R R *>
```

Figure 5.17 Predicted sequence of *A. pisum* leucomyosuppressin prepro-peptide. Sequence data assembled from *A. pisum* genomic clone reads. The sequence shown (from bases 1303 – 3331 of a genomic contig) covers the longest ORF; amino acids shown grey are unlikely to be translated. Probable translated sequence shown as region between brackets < >. Amino acids highlighted in grey form a predicted signal peptide (SignalP v3.0). Amino acids highlighted black comprise the mature leucomyosuppressin peptide. Arrow shows the point of insertion of an intron of approximately 1.6kb in the genomic sequence; intron boundaries are:

- ..CCTATGCCTT^GTGAGTACAT.. (probability of donor site 0.96, NetGene2, *C. elegans* models)
- ..TTTTTCACAG^ACAGCAGAAT.. (probability of acceptor site 1.00, NetGene2, *C. elegans* models)

Bases 106-432 of the sequence shown above are very similar to bases 36-363 of an EST (NCBI identifier CD450787) from brown citrus aphid (*Toxoptera citricida*) which encode a similar polypeptide.

5.5 Discussion

Two strategies for the production of fusion proteins were carried out in this chapter. The first fusion protein contained a venom protein fused to the C-terminus of ASaII. Red scorpion toxin (RST) is a neurotoxin which acts in a target-specific manner to modulate ion channel activity in lepidopteran insects (Wudayagiri, R., *et al.*, 2001). Initial production of full-length ASaII₍₁₂₇₎-RST resulted in proteolytic cleavage of all of the fusion protein produced in *P. pastoris*. Information obtained in this report (Chapter 3.1.9) suggested that this cleavage was occurring at proteinase A site at the C-terminus of ASaII. Subsequently the construct was redesigned to encode only the first 115 amino acids of ASaII, ASaII₍₁₁₅₎-RST. This construct produced intact fusion protein. RST had been shown previously to demonstrate toxicity to *N. lugens* in a fusion protein of GNA-RST (FP4) (Pham Trung, N., *et al.*, 2006). Subsequently it was thought that an ASaII-RST construct may be potentially toxic to *A. pisum*, given that ASaII showed greater toxicity to *A. pisum* than GNA. However, artificial diet bioassays against *A. pisum* showed that at a concentration of 1 mg/ml no toxicity was observed. This could be accounted for either by the inability of ASaII to cross the aphid gut or a conformational problem in the intact fusion protein.

The use of peptides in fusion proteins provides a potentially highly specific approach to insect control, in this study biotinylated leucomyosuppressin (LMS) from cockroach, *D. punctata* was used as a potential toxin in a fusion protein ASaII-Avidin-LMS. By utilizing the avidin-biotin interaction an ASaII-Avidin fusion can be used to conjugate a large number of chemically-synthesised biotinylated peptides for assessment of their toxicity. Production of intact ASaII-Avidin was achieved in *P. pastoris* after several different purification strategies were attempted. Affinity chromatography using the polyhistidine vector tag was established as the most efficient means of purification but in standard, neutral pH buffers 'salting out' of the protein solution occurred. By maintaining the pH at 4.0 it was possible to prevent precipitation of the protein solution and ASaII-Avidin was able to be eluted in standard elution buffer at pH 7.4. Upon elution it was crucial that dialysis was carried out immediately and that the eluate was diluted by

approximately 5-fold to ensure that precipitation of the recombinant protein did not occur during dialysis. After dialysis the protein solution had to be transferred immediately to – 20 °C to prevent proteolysis. Again, this proteolysis may have been occurring at the ASaII C-terminus proteinase A site. Intact ASaII-Avidin was conjugated to LMS and then separation was carried out using gel-filtration chromatography. Intact ASaII-Avidin-LMS was verified and quantified by Western blotting and ELISA for use in artificial diet bioassays against *A. pisum*. Artificial diet bioassays showed ASaII-Avidin-LMS to be toxic to *A. pisum* over the first chronic phase of toxicity (days 4-13). The LMS peptide appears to be acting as an inhibitor of aphid feeding as the expected effect of progressive inhibition of feeding and reduction in honeydew production was observed during the feeding trial. For LMS to exert these anti-feedant effects then *A. pisum* must possess an endogenous LMS peptide. No such peptide has been identified, the only LMS peptide identified in hemipteran insects to date has been in the brown citrus aphid, *Toxoptera citricida* (NCBI identifier CD450787). Using the sequence from *T. citricida* it was possible to deduce a putative *A. pisum* LMS prepro-peptide sequence using sequence data from genomic clones. The putative sequence contains a predicted signal peptide and a mature QDLdHVFLRF peptide sequence. The high level of homology between LMS sequences means that the effects of the endogenous *A. pisum* LMS peptide will be mimicked and modulated to a pseudo-constitutive level by the presence of ASaII-Avidin-LMS in the haemolymph of *A. pisum*.

6 General Discussion

Hemipteran insects represent a significant order of insect pest species responsible for large amounts of crop damage across the world. These insects are also responsible for the transmission of many plant viruses that in turn result in further crop damage. The problem of insect pest damage will become increasingly important in UK agriculture as the effects of global climate change manifest themselves. In particular strategies to deal with hemipteran pests need to be developed with species such as *Acyrtosiphon pisum* being described as a moderate pest of pea plants and *Myzus persicae* being described as a major pest of potatoes, lettuces, legumes and brassicas by the Rothamstead Insect Survey. Both of these species are carriers of plant viruses and aphids in general are particularly likely to be able to respond to changes in climate due to their short generation time and rapid growth rate, there is evidence to suggest that increases in temperature with increase the speed of aphid development (Bayhan, E., *et al.* (2005). Additionally, conventional insecticides are becoming less effective against some hemipteran pests as resistant populations emerge (Koss, A. M. *et al.* 2005). Genetic modification represents a potential solution to the problem of insect pest damage, however at present this strategy is not in favour in the UK and established lines containing *Bacillus thuringiensis* toxin (*Bt* toxin) do not provide resistance to hemipteran insects. The use of lectins provides an alternative to *Bt* toxin, GNA has been shown to have limited toxicity towards *N. lugens* (Down, R. E., *et al.*, 2006) and a lectin extracted from garlic leaves has been shown to have toxicity towards *L. erysimi* (Banerjee, S., *et al.*, 2004). Transgenic plants expressing the garlic leaf lectin (ASAL) have been produced; ASAL expressing tobacco significantly reduce the survival of *M. persicae* compared to wild-type plants (Dutta, I., *et al.*, 2005). Survival of rice brown planthopper (*N. lugens*) and green leafhopper (*Nephotettix virescens*) was also significantly reduced on transgenic rice plants expressing ASAL. Subsequently the use of garlic bulb lectins produced in a recombinant yeast expression system was investigated as a potential insect control agent being a protein which is consumed in large quantities by humans without harmful effects.

symbionin, a protein secreted from *B. aphidicola* an endosymbiotic bacteria was important in the mode of toxicity to *L. erysimi*. This implied that aposymbiotic insects may show a reduced susceptibility to garlic bulb lectins. To investigate the significance of the bacterial endosymbiont on insect survival, artificial diet bioassays against symbiotic and aposymbiotic *A. pisum* and *M. persicae* were carried out. No significant difference could be determined between symbiotic and aposymbiotic insects, suggesting the assertion that the mechanism of toxicity involves a symbiont-produced protein is incorrect. Banerjee *et al.* observed the interaction between symbionin and garlic leaf lectin by using a far-Western blotting technique, in this study lectin-gut interactions were studied in *A. pisum* using a 'ligand-fishing' approach utilizing ASaII conjugated magnetic beads. This approach showed no evidence of an interaction between symbionin and ASaII. Two proteins were identified from prominent bands on SDS-PAGE using MALDI-TOF spectroscopy: alanyl aminopeptidase N (APN) and sucrase. APN represents a major constituent of the *A. pisum* gut, accounting for over 15% of total gut proteins. *A. pisum* APN resembles APN seen in lepidopteran insects, containing standard residues involved in catalysis and a glycosyl-phosphatidylinositol anchor (GPI). APN is a target for *Bt* toxins in lepidopteran insects (Knight, P. J., *et al.*, 1995), (Valaitis, A. P., *et al.*, 1995) and contains large amounts of mannose oligomers which explains its ability to bind other entomotoxic lectins such as Con A. The APN from *A. pisum* has a wide specificity to N-terminal amino acids and has been shown to prefer tripeptides (Cristofaletti, P. T., *et al.*, 2006). Saturation of APN enzymes in the aphid gut by ASaII could eliminate its activity and lead to insect death. Despite relatively low levels of proteolytic activity in the aphid gut, a disruption of the activity of APN could be significant in affecting aphid survival. The second protein pulled-down by ASaII was gut sucrase, an α -glucosidase enzyme present in the distal midgut region responsible for the hydrolysis of phloem-derived sucrose into glucose, a sugar which can be assimilated by aphids. The role of sucrase is important in osmoregulatory balance in the aphid (Price, D. R., *et al.*, 2007). It is unclear how the interaction between ASaII and sucrase may occur as sucrase is a non-glycosylated protein, so an interaction between ASaII and mannose residues on the surface of the enzyme can be ruled out, suggesting a protein-protein interaction must be occurring. Evidence of a lack of sucrase activity was observed in

insects fed diet containing acarbose (an α -glucosidase inhibitor), the resultant insects were seen to be 'shrivelled' (Price, D. R., *et al.*, 2007), indicating the movement of water from the body fluid into the gut. Prior to death the aphids fed on recombinant lectins appeared to be slightly shrivelled although it is impossible to assert that this was an effect of reduced sucrase activity. If the toxicity of recombinant garlic lectins occurs via their ability to bind APN and sucrase in the aphid gut and remove the function of these two enzymes then it would be expected that large amounts of the lectin would be needed to produce a toxic effect and this toxic effect would not be acute. This concurs with what is seen in the insect bioassay data presented in this study where insect death occurred predominately in the first stage of chronic toxicity between days 4-10.

Another property of plant lectins is their resistance to proteolysis and stability within the insect gut. Many lectins, such as GNA are able to cross the insect gut and be detected in the insect haemolymph as intact proteins (Fitches, E., *et al.*, 2001). This property is of significance as lectins can be used as carrier molecules to transport insecticidal proteins and peptides into the insect haemolymph, making lectins a valuable tool for the oral administration of toxins. A number of these fusion protein constructs have been produced and been shown to have significantly higher toxicity to insect pest species when compared to the lectin or toxin subunits alone (Fitches, E., *et al.*, 2002, Fitches, E., *et al.*, 2004). Recombinant garlic lectin offers an alternative for use in fusion protein constructs with potentially greater transepithelial transport ability. However, it was not possible in this study to determine whether or not recombinant ASaII was transported into the aphid haemolymph. The sensitivities of both the vector tag antibody and an antibody raised against an immunogenic region of ASaII were too low to allow detection of ASaII in the haemolymph, to confirm the presence or absence of ASaII transport an antibody with greater sensitivity is required. During the course of this work three attempts to raise antibodies to recombinant garlic lectins were made, none of which resulted in an antibody that was immunoreactive at useful concentrations (<20 ng) for the detection of recombinant lectins in the haemolymph. Using anti-avidin antibodies it was possible to detect the avidin subunit of an ASaII-avidin fusion; however, there was no evidence of intact ASaII-avidin. This suggested that avidin alone may be transported into the

haemolymph, a result that was confirmed in this study using aphids fed on diet containing avidin only. Work in this study showed that an ASaII-avidin-LMS fusion protein had greater toxicity to *A. pisum* than both ASaII-avidin and LMS at equivalent concentrations. Further work would involve analysing variants of the ASaII-avidin-LMS fusion construct, for example comparing the relative toxicities of ASaII-LMS against avidin-LMS, and comparing this data with data relating to transport of ASaII and avidin. When considering binding and transport of lectins across the insect gut little is known about what properties enhance or reduce the ability of a lectin to be transported into the insect haemolymph, in the case of ASaII strong binding to surface carbohydrates on APN may inhibit its ability to cross the insect gut, whereas lectins with less stringent interactions may be sequestered at the extracellular gut surface before being more readily released from these interactions and transported across the gut into the haemolymph. At present very little is known about the mechanisms of transepithelial transport of proteins across the insect gut and reports that have been published have suggested different mechanisms. Subsequently there is no precedent for evaluating a general model of protein transport across the insect gut.

The toxicity of ASA-avidin-LMS is the first successful demonstration of the use of a chemically synthesized peptide in a fusion protein construct. The use of avidin in the construct provides a convenient means of coupling a wide range of biotinylated peptides for assessment of their toxicity, which eliminates the need for subcloning and expression of fusion proteins containing small peptides which are difficult to produce. The toxicity of the LMS peptide from *D. punctata* against *A. pisum* suggested the presence of an endogenous LMS, the activity of which was being mimicked by the presence of the exogenous LMS. Aphids fed ASA-avidin-LMS were reduced in size compared to control insects and produced lower levels of honeydew, which may be accounted for by a breakdown in normal gut peristaltic movements due to the effect of LMS. Using genomic DNA data from the *A. pisum* expressed sequence tag (EST) database it was possible to assemble a gene sequence for a putative LMS in *A. pisum*.

The lack of toxicity of ASaII₍₁₂₇₎-RST was initially thought to be due to proteolytic cleavage of the fusion protein resulting in the insecticidal RST component not being delivered to the haemolymph. Redesign of this construct to make ASaII₍₁₁₅₎-RST was carried out to keep the fusion protein intact, this operation was successful but a toxic effect was still not observed when *A. pisum* bioassays were carried out. A fusion protein of GNA-RST (FP4) was shown to be toxic to *N. lugens*, providing evidence that RST, which was previously thought to be a lepidopteran-specific toxin, was toxic to at least one hemipteran pest. Bioassays showed ASaII₍₁₁₅₎-RST is not toxic to *A. pisum*, however ASaII₍₁₁₅₎-RST also failed to demonstrate toxicity to tomato moth, *L. oleracea* (E. C. Fitches, personal communication).

One of the aims of the project was to investigate whether or not recombinant garlic lectins may offer a viable alternative carrier protein for incorporation into fusion proteins. It was not possible to analyse the ability of ASaII to cross the insect gut into the haemolymph which has made direct comparison between the transport of GNA and ASaII impossible. The results of artificial diet bioassays of ASaII-RST against *A. pisum* have demonstrated that it is non-toxic. Similar results have been observed in *L. oleracea*, which suggests that in the case of RST at least, a similar fusion protein to GNA-RST using ASaII is less toxic.

In conclusion production of recombinant garlic bulb lectins in a yeast expression system yields functionally active proteins with agglutination activity and toxicity to a number of hemipteran insect pests. The interaction of a recombinant garlic bulb lectin, ASaII, at the gut surface in *A. pisum* is with the widespread gut enzyme APN and also with membrane-bound gut sucrase. The mode of toxicity may be to render APN and gut sucrase inactive, causing insect death within the first chronic toxicity phase. The ability of ASaII to cross the *A. pisum* gut remains unclear, although it has been shown that ASaII-avidin may represent an effective carrier of small peptides such as LMS from cockroach (*D. punctata*), resulting in increased toxicity compared to ASaII-avidin and LMS alone. Using *A. pisum* genomic clones it was possible to assemble a putative *A. pisum* LMS gene which is the basis for the mode of toxicity of ASaII-avidin-LMS.

7 References

Angata, T. and Brinkman-Van Der Linden, E. C. M. I-Type lectins. *Biochimica et Biophysica Acta*. (1572), 294-316, (2002).

Bachhawat, K., Kapoor, M., Dam, T. K. and Suriola, A. The Reversible Two-State Unfolding of a monocot Mannose-Binding Lectin from Garlic Bulbs Reveals the Dominant Role of the Dimeric Interface in Its Stabilization. *Biochemistry*. (40), 7291-7300, (2001).

Balzarini, J., Schols, D., Neyts, J., Van Damme, E. J., Peumans, W. J. and De Clerq, E. α -(1, 3) and α -(1, 6)-mannose specific plant lectins are markedly inhibitory to human immunodeficiency virus and cytomegalovirus infections *in vitro*. *Antimicrobial Agents Chemotherapy*. (35), 410-416, (1991).

Bandyopadhyay, S., Roy, A. and Das, S. Binding of garlic (*Allium sativum*) leaf lectin to the gut receptors of homopteran pests is correlated to its insecticidal activity. *Plant Science*. (161), 1025-1033, (2001).

Banerjee, S., Hess, D., Majumder, P., Roy, D. and Das, S. The interactions of *Allium sativum* leaf agglutinin with a chaperonin group of unique receptor protein Isolated from a bacterial endosymbiont of the mustard aphid. *J Biol Chem*. 279(22), 23782-23789. (2004).

Baumgartner, P., Harper, K., Raemaekers, R. J., Durieux, A., Gatehouse, A. M., Davies, H. V. and Taylor, M. A. Large-scale production and purification of recombinant *Galanthus nivalis* agglutinin (GNA) expressed in the methylotrophic yeast *Pichia pastoris*. *Biotechnol Lett*. (25), 1281-5, (2003).

- Baumgartner, P., Raemaekers, R. J., Durieux, A., Gatehouse, A., Davies, H. and Taylor, M. Large-scale production, purification, and characterisation of recombinant *Phaseolus vulgaris* phytohemagglutinin E-form expressed in the methylotrophic yeast *Pichia pastoris*. *Protein Expr Purif.* (26), 394-405, (2002).
- Bayhan, E., Olmez-Bayhan, S., Ulusoy, M. R. and Brown, J. K. Effect of temperature on the biology of *Aphis punicae* (Passerini) (Homoptera: Aphididae) on pomegranate. *Environmental Entomology*, 34(1), 22-26, (2005).
- Beintema, J. J. Structural features of plant chitinases and chitin-binding proteins. *FEBS Lett.* (350), 159-63, (1994).
- Bell, H. A., Fitches, E. C., Down, R. E., Marris, G. C., Edwards, J. P., Gatehouse, J. A. and Gatehouse, A. M. The effect of snowdrop lectin (GNA) delivered via artificial diet and transgenic plants on *Eulophus pennicornis* (Hymenoptera: Eulophidae), a parasitoid of the tomato moth *Lacanobia oleracea* (Lepidoptera: Noctuidae). *J Insect Physiol.* (45), 983-991, (1999).
- Bennett, M. J., Choe, S. and Eisenberg, D. Domain swapping: Entangling alliances between proteins. *Proceedings of the National Academy of Sciences of the United States of America.* (91), 3127-3131, (1994).
- Beyenbach, K. W. Transport mechanisms of diuresis in Malpighian tubules of insects. *J Exp Biol.* (206), 3845-56, (2003).
- Blum, H., Beier, H. and Gross H. J. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis.* 8 (2), 93-99 (1987).
- Borovsky, D. and Mahmood, F. Feeding the mosquito *Aedes aegypti* with TMOF and its analogs; effect on trypsin biosynthesis and egg development. *Regul Pept.* (57), 273-81, (1995).

Bowles, D. J., Marcus, S. E., Pappin, D. J., Findlay, J. B., Eliopoulos, E., Maycox, P. R. and Burgess, J. Posttranslational processing of concanavalin A precursors in jackbean cotyledons. *J Cell Biol.* (102), 1284-97, (1986).

Boyd, W. C. and Shapleigh, E. Separation of Individuals of Any Blood Group into Secretors and Non-Secretors by Use of a Plant Agglutinin (Lectin). *Blood.* (9), 1195-1198, (1954).

Boyko, E. V., Smith, C. M., Thara, V. K., Bruno, J. M. and Deng, Y. The molecular basis of plant gene expression during aphid invasion: wheat Pto- and Pti-like sequences are involved in interactions between wheat and Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology.* (99), 1430-1445, (2006).

Casartelli, M., Corti, P., Giovanna Leonardi, M., Fiandra, L., Burlini, N., Pennacchio, F. and Giordana, B. Absorption of albumin by the midgut of a lepidopteran larva. *J Insect Physiol.* (51), 933-40, (2005).

Chandra, N. R., Ramachandraiah, G., Bachhawat, K., Dam, T. K., Surolia, A. and Vijayan, M. Crystal structure of a dimeric mannose-specific agglutinin from garlic: quaternary association and carbohydrate specificity. *J Mol Biol.* (285), 1157-68, (1999).

Chao, Q., Casalongue, C., Quinn, J. M. and Etzler, M. E. Expression and partial characterization of *Dolichos biflorus* seed lectin in *Escherichia coli*. *Arch Biochem Biophys.* (313), 346-50, (1994).

Cherqui, A. and Tjallingii, W. F. Salivary Proteins of Aphids, a Pilot Study on Identification, Separation and Immunolocalisation. *Journal of Insect Physiology.* (46), 1177-1186, (2000).

Cohen, S. L. and Chait, B. T. Mass Spectrometry of Whole Proteins Eluted from Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis Gels. *Analytical Biochemistry*. (247), 257-267, (1997).

Conde, R., Cueva, R., Pablo, G., Polaina, J. and Larriba, G. A search for hyperglycosylation signals in yeast glycoproteins. *J Biol Chem*. (279), 43789-98, (2004).

Cristofolletti, P. T., De Sousa, F. A., Rahbe, Y. and Terra, W. R. Characterization of a membrane-bound aminopeptidase purified from *Acyrtosiphon pisum* midgut cells. A major binding site for toxic mannose lectins. *Febs J*. (273), 5574-88, (2006).

Dahms, N. M. and Hancock, M. K. P-Type lectins. *Biochimica et Biophysica Acta*. (1572), 317-340, (2002).

Dam, T. K., Bachhawat, P. K., Rani, G. and Suriola, A. Garlic (*Allium sativum*) Lectins Bind to high Mannose Oligosaccharide Chains. *Journal of Biological Chemistry*. (273), 5528-5535, (1998).

Dodd, R. B. and Drickamer, K. Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity. *Glycobiology*. (11), 71R-79R, (2001).

Drickamer, K. C-type lectin domains. *Current Opinions in Structural Biology*. (9), 585-590, (1999).

Donly, B. C., Fuse, M., Orchard, I., Tobe, S. S. and Bendena, W. G. Characterization of the gene for leucomyosuppressin and its expression in the brain of the cockroach *Diploptera punctata*. *Insect Biochem Mol Biol*. (26), 627-37, (1996).

Douglas, A. E., Minto, L. B. and Wilkinson, T. L. Quantifying nutrient production by the microbial symbionts in an aphid. *Journal of Experimental Biology*. (204), 349-358, (2001).

Down, R. E., Fitches, E. C., Wiles, D. P., Corti, P., Bell, H. A., Gatehouse, J. A. and Edwards, J. P. Insecticidal spider venom toxin fused to snowdrop lectin is toxic to the peach-potato aphid, *Myzus persicae* (Hemiptera: Aphididae) and the rice brown planthopper, *Nilaparvata lugens* (Hemiptera: Delphacidae). *Pest Manag Sci.* (62), 77-85, (2006).

Dreyer, T. Substrate specificity of proteinase yscA from *Saccharomyces cerevisiae*. *Carlsberg Res Commun.* (54), 85-97, (1989).

Du, J., Foissac, X., Carss, A., Gatehouse, A. M. and Gatehouse, J. A. Ferritin acts as the most abundant binding protein for snowdrop lectin in the midgut of rice brown planthoppers (*Nilaparvata lugens*). *Insect Biochem Mol Biol.* (30), 297-305, (2000).

Dutta, I., Saha, P., Majumder, P., Sarkar, A., Chakraborti, D., Banerjee, S. and Das, S. The efficacy of a novel insecticidal protein, *Allium sativum* leaf lectin (ASAL), against homopteran insects monitored in transgenic tobacco. *Plant Biotechnol J.* (3), 601-11, (2005).

Fiedler, K. and Simons, K. A putative novel class of animal lectins in the secretory pathway homologous to leguminous lectins. *Cell.* (77), 625-626, (1994).

Fishman, L. and Zlotkin, E. A diffusional route of transport of horseradish peroxidase through the midgut of a fleshfly. *Journal of Experimental Zoology.* (229), 189-195, (1984).

Fitches, E., Audsley, N., Gatehouse, J. A. and Edwards, J. P. Fusion proteins containing neuropeptides as novel insect control agents: snowdrop lectin delivers fused allatostatin to insect haemolymph following oral ingestion. *Insect Biochem Mol Biol.* (32), 1653-61, (2002).

Fitches, E., Edwards, M. G., Mee, C., Grishin, E., Gatehouse, A. M., Edwards, J. P. and Gatehouse, J. A. Fusion proteins containing insect-specific toxins as pest control agents: snowdrop lectin delivers fused insecticidal spider venom toxin to insect haemolymph following oral ingestion. *J Insect Physiol.* (50), 61-71, (2004a).

Fitches, E., Wilkinson, H., Bell, H., Bown, D. P., Gatehouse, J. A. and Edwards, J. P. Cloning, expression and functional characterisation of chitinase from larvae of tomato moth (*Lacanobia oleracea*): a demonstration of the insecticidal activity of insect chitinase. *Insect Biochem Mol Biol.* (34), 1037-50, (2004b).

Fitches, E., Woodhouse, S. D., Edwards, J. P. and Gatehouse, J. A. In vitro and in vivo binding of snowdrop (*Galanthus nivalis* agglutinin; GNA) and jackbean (*Canavalia ensiformis*; Con A) lectins within tomato moth (*Lacanobia oleracea*) larvae; mechanisms of insecticidal action. *J Insect Physiol.* (47), 777-787, (2001).

Foissac, X., Edwards, M. G., Du, J. P., Gatehouse, A. M. R. and Gatehouse, J. A. Putative protein digestion in a sap-sucking homopteran plant pest (rice brown plant hopper; *Nilaparvata lugens*: Delphacidae)- identification of trypsin-like and cathepsin B-like proteases. *Insect Biochemistry Molecular Biology.* (32), 967-978, (2002).

Fonagy, A., Schoofs, L., Proost, P., Van Damme, J., Bueds, H. and De Loof, A. Isolation, primary structure and synthesis of neomyosuppressin, a myoinhibiting neuropeptide from the grey fleshfly, *Neobellieria bullata*. *Comp Biochem Physiol C.* (102), 239-45, (1992).

Gatehouse, J. A., A, M. R. G. and Fitches, E. Effects of snowdrop lectin (GNA) delivered via artificial diet and transgenic plants on the development of tomato moth (*Lacanobia oleracea*) larvae in laboratory and glasshouse trials. *J Insect Physiol.* (43), 727-739, (1997).

Gill, S. S., Cowles, E. A. and Francis, V. Identification, isolation and cloning of a *Bacillus thuringiensis* Cry1Ac toxin-binding protein from the midgut of the lepidopteran

insect *Heliothis virescens*. Journal of Biological Chemistry. 270 (45) 27277-27282, (1995).

Habibi, J., Brandt, S. L., Coudron, T. A., Wagner, R. M., Wright, M. K., Backus, E. A. and Huesing, J. E. Uptake, flow, and digestion of casein and green fluorescent protein in the digestive system of *Lygus hesperus* Knight. Arch Insect Biochem Physiol. (50), 62-74, (2002).

Hester, G., Kaku, H., Goldstein, I. J. and Wright, C. S. Structure of mannose-specific snowdrop (*Galanthus nivalis*) lectin is representative of a new plant lectin family. Nat Struct Biol. (2), 472-9, (1995).

Higgins, D. R. and Creggs, J. M., *Pichia* protocols. Methods in Molecular Biology, 1998, Human Press.

Holman, G. M., Cook, B. J. and Nachman, R. J. Primary structure and synthesis of a blocked myotropic neuropeptide isolated from the cockroach, *Leucophaea maderae*. Comp Biochem Physiol C. (85), 219-24, (1986).

Ishiwatari, Y., Honda, C., Kawashima, I., Nakamura, S., Hirano, H., Mori, S., Fugiwara, T., Hayashi, H. and M., C. Thioredoxin-H is one of the Major Proteins in Rice Phloem Sap, Planta. (195), 456-463, (1995).

Jeffers, L. A., Thompson, D. M., Ben-Yakir, D. and Roe, R. M. Movement of proteins across the digestive system of the tobacco budworm, *Heliothis virescens*. Entomologia Experimentalis Et Applicata. (117), 135-146, (2005).

Kaku, H., Goldstein, I. J., Van Damme, E. J. and Peumans, W. J. New Mannose-Specific Lectins from Garlic (*Allium sativum*) and Ramsons (*Allium ursinum*). Carbohydrate Research. (229), 347-353, (1992).

Kaku, H., Tanaka, Y., Tazaki, K., Minami, E., Mizuno, H. and Shibuya, N. Sialylated oligosaccharide-specific plant lectin from japanese elderberry (*Sambucus sieboldiana*) bark tissue has a homologous structure to type II ribosome-inactivating proteins, ricin and arbrin. *Journal of Biological Chemistry*. (271), 1480-1485, (1996).

Kaloshian, I. Gene-for-gene disease resistance: bridging insect pest ans pathogen defence. *Journal of Chemical Ecology*. (30), 2419-2438, (2004).

Keitel, U., Simon, O., Borriss, R. and Heinemann, U. Molecular and active site structure of of a *Bacillus* 1,3-1,4-b-glucanase. *Proc Natl Acad Sci U S A*. (90), 5287-5291, (1993).

Kingan, T. G., Teplow, D. B., Phillips, J. M., Riehm, J. P., Rao, K. R., Hildebrand, J. G., Homberg, U., Kammer, A. E., Jardine, I., Griffin, P. R. and Et Al. A new peptide in the FMRFamide family isolated from the CNS of the hawkmoth, *Manduca sexta*. *Peptides*. (11), 849-56, (1990).

Knibbs, R., Goldstein, I. J., Ratcliff, R. M. and Shibuya, N. Characterization of the carbohydrate binding specificity of the leuko-agglutinating lectin from *Maackia anurensis*. *Journal of Biological Chemistry*. (266), 83-88, (1991).

Kohanski, R. A. and Lane, M. D. Monovalent avidin affinity columns. *Methods Enzymol*. (184), 194-200, (1990).

Knight, P. J., Knowles, B. H. and Ellar, D. J. Molecular cloning of an insect aminopeptidase N that serves as a receptor for *Bacillus thuringiensis* CryIA(c) toxin. *J Biol Chem*. (270), 17765-70, (1995).

Koss, A. M., Jensen, A. S., Schreiber, A., Pike, K. S. and Snyder, W. E. Comparison of predator and pest communities in Washington potato fields treated with broad-spectrum, selective and organic insecticides. *Environmental Ecology*, 34(1), 87-95. (2005).

Kuhn, P., Tarentino, A. L., Plummer, T. H. and Van Roey, P. Crystal structure of peptide N(4)-(N-acetyl-b-D-glucosminyl) asparagine amidase at 2.2 Å resolution. *Biochemistry*. (33), 11699-11706, (1994).

Kyogashima, M., Ginsburg, V. and Krivan, H. C. *Escherichia coli* K99 binds to N-glycolyl-sialoparaglobosidase and N-glycolyl-GM3 found in piglet small intestine. *Arch Biochem Biophys*. (270), 391-397, (1989).

Lamb, F. I., Roberts, L. M. and Lord, J. M. Nucleotide sequence of cloned cDNA coding for preproricin. *Eur J Biochem*. (148), 265-270, (1985).

Låemmli, U. K. Cleavage of structural proteins during the assembly of the head of Bacteriophage T4. *Nature*. (227), 680-685, (1970).

Lin, J., Yao, J., Zhou, X., Sun, X. and Tang, K. Expression and purification of a novel mannose-binding lectin from *Pinellia ternata*. *Mol Biotechnol*. (25), 215-22, (2003).

Longstaff, M., Powell, K. S., Gatehouse, J. A., Raemaekers, R., Newell, C. A. and Hamilton, W. D. Production and purification of active snowdrop lectin in *Escherichia coli*. *Eur J Biochem*. (252), 59-65, (1998).

Loris, R. Principles of structures of animal and plant lectins. *Biochim Biophys Acta*. (1572), 198-208, (2002).

Madrell, S. H. P. and Gardiner, B. O. C. Excretion of alkaloids by Malpighian tubules of insects. *Journal of Experimental Biology*. (64), 499-508, (1976).

Meyers, B. C., Kozik, A., Griego, A., Kuang, H. H. and Michelmore, R. W. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell*. (15), 809-834, (2003).

Min, W., Dunn, A. J. and Jones, D. H. Non-glycosylated recombinant pro-concanavalin A is active without polypeptide cleavage. *Embo J.* (11), 1303-7, (1992).

Mitsuhashi, J. Methods for rearing leafhoppers and planthoppers on artificial diets. *Review of Plant Protection Research.* 7, 57-67, (1974).

Modespacher, U. P., Rudin, W., Jenni, L. and Hecker, H. Transport of peroxidase through the midgut epithelium of *Glossinia m. morsitans* (Diptera, Glossinidae). *Tissue & Cell.* (18), 429-436, (1986).

Nayak, P., Basu, D., Das, S., Basu, A., Ghosh, D., Ramakrishnan, N.A., Ghosh, M. and Sen, S. K. Transgenic elite indica rice plants expressing CryIAc delta-endotoxin of *Bacillus thuringiensis* are resistant against yellow stem borer (*Scirpophaga incertulas*). *Proc Natl Acad Sci USA.* 94(6) 2111-6. (1997).

Nichols, R. Isolation and structural characterization of Drosophila TDVDHVFLRFamide and FMRFamide-containing neural peptides. *J Mol Neurosci.* (3), 213-8, (1992).

Ofek, I., Mirelman, D. and Sharon, N. Adherence of *Escherichia coli* to human mucosal cells mediated by mannose receptors. *Nature.* (265), 623-625, (1977).

Peumans, W. J. and Van Damme, E. J. M. Lectins as Plant Defense Proteins. *Plant Physiology.* (109), 347-352, (1995).

Pham Trung, N., Fitches, E. and Gatehouse, J. A. A fusion protein containing a lepidopteran-specific toxin from the South Indian red scorpion (*Mesobuthus tamulus*) and snowdrop lectin shows oral toxicity to target insects. *BMC Biotechnol.* (6), 18, (2006).

Powell, K. S., Gatehouse, A. M. R., Hilder, V. A. and Gatehouse, J. A. Antifeedant Effects of Plant-Lectins and an Enzyme on the Adult Stage of the Rice Brown

-
- Planthopper, *Nilaparvata lugens*. Entomologia Experimentalis Et Applicata. (75), 51-59, (1995).
- Powell, K. S., Gatehouse, A. M. R., Hilder, V. A., Vandamme, E. J. M., Peumans, W. J., Boonjawat, J., Horsham, K. and Gatehouse, J. A. Different Antimetabolic Effects of Related Lectins Towards Nymphal Stages of *Nilaparvata Lugens*. Entomologia Experimentalis Et Applicata. (75), 61-65, (1995).
- Price, D. R., Karley, A. J., Ashford, D. A., Isaacs, H. V., Pownall, M. E., Wilkinson, H. S., Gatehouse, J. A. and Douglas, A. E. Molecular characterisation of a candidate gut sucrase in the pea aphid, *Acyrtosiphon pisum*. Insect Biochem Mol Biol. (37), 307-17, (2007).
- Prosser, W. A. and Douglas, A. E. A test of the hypotheses that nitrogen is upgraded and recycled in an aphid (*Acyrtosiphon pisum*) symbiosis. Journal of Insect Physiology. (33)2, 93-99, (1992).
- Raemaekers, R. J., De Muro, L., Gatehouse, J. A. and Fordham-Skelton, A. P. Functional phytohemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris* correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. Eur J Biochem. (265), 394-403, (1999).
- Rahbe, Y., Sauvion, N., Febvay, G., Peumans, W. J. and Gatehouse, A. M. R. Toxicity of Lectins and Processing of Ingested Proteins in the Pea Aphid *Acyrtosiphon pisum*. Entomologia Experimentalis Et Applicata. (76), 143-155, (1995).
- Raikhel, N. V. and Wilkins, T. A. Isolation and Characterization of a Cdna Clone Encoding Wheat-Germ-Agglutinin. Proceedings of the National Academy of Sciences of the United States of America. (84), 6745-6749, (1987).

-
- Rao, K. V., Rathore, K. S., Hodges, T. K., Fu, X., Stoger, E., Sudhakar, D., Williams, S., Christou, P., Bharathi, M., Bown, D. P., Powell, K. S., Spence, J., Gatehouse, A. M. and Gatehouse, J. A. Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper. *Plant J.* (15), 469-77, (1998).
- Rigas, D. A. Purification and properties of the phytohemagglutinin of *Phaseolus vulgaris*. *J Biol Chem.* (212), 607-615, (1955).
- Robb, S., Packman, L. C. and Evans, P. D. Isolation, primary structure and bioactivity of schistoflrf-amide, a FMRF-amide-like neuropeptide from the locust, *Schistocerca gregaria*. *Biochem Biophys Res Commun.* (160), 850-6, (1989).
- Rutenber, E. and Robertus, J. D. Structure of ricin B chain at 2.6 Å resolution. *Proteins.* (10), (1991).
- Saha, P., Majumder, P., Dutta, I., Ray, T., Roy, S. C. and Das, S. Transgenic rice expressing *Allium sativum* leaf lectin with enhanced resistance against sap-sucking insect pests. *Planta.* (223), 1329-43, (2006).
- Salit, I. E. and Gotschlich, E. C. Type I *Escherichia coli* pili: characterization of binding to monkey kidney cells. *Journal of Experimental Medicine.* (146), 1182-1194, (1997).
- Sambrook, J. and Russell, D. W., *Molecular Cloning: A Laboratory Manual*, 2001, CSHL Press, 0-87969-577-3.
- Sauvion, N., Charles, H., Febvay, G. and Rahbe, Y. Effects of jackbean lectin (ConA) on the feeding behaviour and kinetics of intoxication of the pea aphid, *Acyrtosiphon pisum*. *Entomologia Experimentalis Et Applicata.* (110), 31-44, (2004).
- Sauvion, N., Rahbe, Y., Peumans, W. J., Vandamme, E. J. M., Gatehouse, J. A. and Gatehouse, A. M. R. Effects of GNA and other mannose binding lectins on development

and fecundity of the peach-potato aphid *Myzus persicae*. *Entomologia Experimentalis Et Applicata*. (79), 285-293, (1996).

Schoofs, L., Holman, G. M., Paemen, L., Veelaert, D., Amelinckx, M. and De Loof, A. Isolation, identification, and synthesis of PDVDHFLRFamide (SchistoFLRFamide) in *Locusta migratoria* and its association with the male accessory glands, the salivary glands, the heart, and the oviduct. *Peptides*. (14), 409-21, (1993).

Sheldon, P. S. and Bowles, D. J. The glycoprotein precursor of concanavalin A is converted to an active lectin by deglycosylation. *Embo J*. (11), 1297-301, (1992).

Sheldon, P. S., Keen, J. N. and Bowles, D. J. Post-translational peptide bond formation during concanavalin A processing in vitro. *Biochem J*. (320 (Pt 3)), 865-70, (1996).

Smith, C. M. and Boyko, E. V. The molecular bases of plant resistance and defense responses to aphid feeding: current status. *Entomologia Experimentalis Et Applicata*. (122), 1-16, (2006).

Stubbs, M. E., Carver, J. P. and Dunn, R. J. Production of pea lectin in *Escherichia coli*. *J Biol Chem*. (261), 6141-4, (1986).

Tateno, H., Winter, H. C. and Goldstein, I. J. Cloning, expression in *Escherichia coli* and characterization of the recombinant Neu5Acalpha2,6Galbeta1,4GlcNAc-specific high-affinity lectin and its mutants from the mushroom *Polyporus squamosus*. *Biochem J*. (382), 667-75, (2004).

Tu, J., Zhang, G., Datta, K., Xu, C., He, Y., Zhang, Q., Khush, G. S. and Datta, S. K. Field performance of transgenic elite commercial hybrid rice expressing *Bacillus thuringiensis* delta-endotoxin. *Nat. Biotech*. 18(10) 1101-4, (2000).

Valaitis, A. P., Lee, M. K., Rajamohan, F. and Dean, D. H. Brush border membrane aminopeptidase-N in the midgut of the gypsy moth serves as the receptor for the CryIA(c) delta-endotoxin of *Bacillus thuringiensis*. *Insect Biochem Mol Biol.* (25), 1143-51, (1995).

Van Damme, E. J. M., Balzarini, J., Smeets, K., Vanleuven, F. and Peumans, W. J. The Monomeric and Dimeric Mannose-Binding Proteins from the Orchidaceae Species *Listera-Ovata* and *Epipactis Helleborine* - Sequence Homologies and Differences in Biological-Activities. *Glycoconjugate Journal.* (11), 321-332, (1994).

Van Damme, E. J., Barre, A., Rouge, P., Van Leuven, F. and Peumans, W. J. Characterization and molecular cloning of SNAV (nigrin b), a GalNAc-specific type 2 ribosome-inactivating protein from the bark of elderberry (*Sambucus nigra*). *Eur J Biochem.* (237), 505-513, (1996a).

Van Damme, E. J., Barre, A., Rouge, P., Van Leuven, F. and Peumans, W. J. The NeuAca-2,6-gal/galNAc binding lectin from elderberry (*Sambucua nigra*) bark, a type 2 ribosome-inactivating protein with unusual specificity and structure. *Eur J Biochem.* (235), 128-137, (1996b).

Van Damme, E. J., Culerrier, R., Barre, A., Alvarez, R., Rouge, P. and Peumans, W. J. A Novel Family of Lectins Evolutionarily Related to Class V Chitinases: An Example of Neofuncitonalization in Legumes. *Plant Physiology.* (144), 662-672, (2007).

Van Damme, E. J., Goldstein, I. J. and Peumans, W. J. A Comparative Study of Mannose-Binding Lectins from the Amaryllidaceae and Alliaceae. *Phytochemistry.* (30), 509-514, (1991a).

Van Damme, E. J., Kaku, H., Perini, F., Goldstein, I. J., Peeters, B., Yagi, F., Decock, B. and Peumans, W. J. Biosynthesis, primary structure and molecular cloning of snowdrop (*Galanthus nivalis* L.) lectin. *Eur J Biochem.* (202), 23-30, (1991b).

Van Damme, E. J., Peumans, W. J., Barre, A. and Rouge, P. Plant lectins: a composite of several distinct families of structurally and evolutionarily related proteins with diverse biological roles. *Critical Review of Plant Sciences*. (17), 575-592, (1998).

Van Damme, E. J. M., Peumans, W. J., Pusztai, A. and Bardocz, S., *Handbook of Plant Lectins: Properties and Biochemical Applications*, 1998, Wiley, 0-41-96445-X.

Van Damme, E. J. M., Smeets, K., Torrekens, S., Vanleuven, F., Goldstein, I. J. and Peumans, W. J. The Closely Related Homomeric and Heterodimeric Mannose-Binding Lectins from Garlic Are Encoded by One-Domain and 2-Domain Lectin Genes, Respectively. *European Journal of Biochemistry*. (206), 413-420, (1992).

Van Damme, E. J. M., Allen, A. K. and Peumans, W. J. Isolation and Characterization of a Lectin with Exclusive Specificity Towards Mannose from Snowdrop (*Galanthus nivalis*) Bulbs. *Febs Letters*. (215), 140-144, (1987).

Van Damme, E. J. M., Balzarini, J., Smeets, K., Vanleuven, F. and Peumans, W. J. The Monomeric and Dimeric Mannose-Binding Proteins from the Orchidaceae Species *Listera ovata* and *Epipactis helleborine* - Sequence Homologies and Differences in Biological-Activities. *Glycoconjugate Journal*. (11), 321-332, (1994).

Wang, Z. Y., Sun, X. F., Wang, F., Tang, K. X. and Zhang, J. R. Enhanced resistance of snowdrop lectin (*Galanthus nivalis* L. agglutinin)-expressing maize to Asian corn borer (*Ostrinia furnacalis* Guenee). *Journal of Integrative Plant Biology*. (47), 873-880, (2005a).

Wang, Z. Y., Zhang, K. W., Sun, X. F., Tang, K. X. and Zhang, J. R. Enhancement of resistance to aphids by introducing the snowdrop lectin gene gna into maize plants. *Journal of Biosciences*. (30), 627-638, (2005b).

Winther, J. R., Kielland-Brandt, M. C., Sorensen, S. O., Svendsen, I. B. and Crunn, A. W. A high-affinity inhibitor of yeast carboxypeptidase Y is encoded by *TFS1* and shows homology to family of lipid-binding proteins. *Biochemistry*. 37(10), 3351-3357, (1998).

Wood, K. A., Lord, J. M., Wawrzynczak, E. J. and Piatak, M. Preproarbrin: genomic cloning, characterization and the expression of the A-chain in *Escherichia coli*. *Eur J Biochem*. (198), 723-732, (1991).

Wudayagiri, R., Inceoglu, B., Herrmann, R., Derbel, M., Choudary, P. V. and Hammock, B. D. Isolation and characterization of a novel lepidopteran-selective toxin from the venom of South Indian red scorpion, *Mesobuthus tamulus*. *BMC Biochem*. (2), 16, (2001).

Yaoi, K., Nakanashi, K., Kadotani, T., Imamura, M., Koizumi, N., Iwahama, H. and Sato, R. cDNA cloning and expression of *Bacillus thuringiensis* Cry1Aa toxin binding 120 kDa aminopeptidase N from *Bombyx mori*. *Biochim Biophys Acta*. 1444 (1) 131-137, (1999).

Zhu, W., Vandingenen, A., Huybrechts, R., Vercammen, T., Baggerman, G., De Loof, A., Poulos, C. P., Velentza, A. and Breuer, M. Proteolytic breakdown of the Neb-trypsin modulating oostatic factor (Neb-TMOF) in the hemolymph of different insects and its gut epithelial transport. *J Insect Physiol*. (47), 1235-1242, (2001).